0.88 0.89 0.89 1.03

160.05 155.05 155.05 169.99 169.99

```
0.0722 \\ 0.1024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0.7585
0.7596
0.7606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     .7609
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              7089
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                7709
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0.5080
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          0.5260
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0.5385
                                                                                                                                                                                                                                                                   175.53
176.99
170.85
170.85
170.85
170.85
170.65
170.65
170.65
170.65
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
170.95
17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 664 50 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 Documentation
                                                                                                                                       About: Results were produced by the GenCore software, version 4.5, Copyright (c) 1993-2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Secore Escore Len i Documentati Sequence sequence adquences gequence pembl/AA1992 DAT: AAR21277 - SIDS1/gcgdata/hold-geneseqg-embl/AA1992 DAT: AAR21277 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR41895 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR41895 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR41899 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR41897 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR41896 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR36505 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR36505 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR038601 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR0438601 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR0438601 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR0438600 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR0438600 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR05058 - SIDS1/gcddata/hold-geneseqg-embl/AA2000 DAT: AAR05058 - SIDS1/gcgdata/hold-geneseqg-embl/AA2000 DAT: AAR05058 - SIDS1/gcddata/hold-geneseqg-
                     out_format : pfs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1996_DAT:AAW05136 |
| SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1996_DAT:AAW05138 |
| SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1996_DAT:AAW05139 |
| SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1996_DAT:AAW05140 |
| SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1996_DAT:AAW05140 |
| SIDSI/gcgdata/hold-geneseq/geneseqp-embl/AA1996_DAT:AAW05140 |
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /SIDS1/gcgdata/hold-geneseg/genesegp-embl/AA1996.DAT:AAW05144
/SIDS1/gcgdata/hold-geneseg/genesegp-embl/AA1996.DAT:AAW05141
/SIDS1/gcgdata/hold-geneseg/genesegp-embl/AA2000.DAT:AAB24812
OM of: US-09-674-779-4 to: A_Geneseq_032802:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Database sequences: 747574
Database length: 111073796
Search time (sec): 52.040000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database: A_Geneseq_032802:*
                                                                          Date: Jul 30, 2002 4:24 PM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Search information block:
Query: US-09-674-779-4
                                                                                                                                                                                                                                          Command line parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query length: 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 score_list:
```

```
Rev-caspases are cysteine proceases that specifically cleave proteins subunit is N-terminal to a large subunit. A gene delivery vehicle comprising a rev-caspase coding sequence is useful for the treatment of cancer, where the gene delivery vehicle is internalised by tumour cells. The gene delivery vehicle can also be used, to treat autoimmune diseases. Cells transfected with a rev-caspase expressing vector can be used in identification of inhibitors or enhancers of caspase—mediated apoptosis. In vitro translated rev-caspase can be used to identify an inhibitor or enhancers of caspase inhibitors are useful for treating neurodegenerative diseases as well as for inhibiting apoptosis in the heart following myocardial infarction. The present sequence represents the amino acid sequence of rev-caspase-6.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rev-caspase; cysteine protease; zymogen; caspase; cancer; gene therapy; autoimmune disease; caspase-mediated apoptosis; neurodegenerative; tumour cell; myocardial infarction; human.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The invention relates to nucleic acid molecules encoding rev-caspases
             57.50
57.50
57.50
57.00
57.00
                                                                                                                                                                                                    seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:AAY21727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New isolated nucleic acid molecule encoding a rev-caspase - used for screening and identifying inhibitors or enhancers for treating
/SIDS1/gcgdata/hold-geneseg/genesegp-embl/AA2000.DAT:AAG44772
/SIDS1/gcgdata/hold-geneseg/genesegp-embl/AA2000.DAT:AAB24811
/SIDS1/gcgdata/hold-geneseg/genesegp-embl/AA2000.DAT:AAG44771
/SIDS1/gcgdata/hold-geneseg/genesegp-embl/AA2001.DAT:ABB41899
/SIDS1/gcgdata/hold-geneseg/genesegp-embl/AA2001.DAT:ABB41899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60 CATGTCTTACCAAAAAAAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Identity: 58.824
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Align seg 1/1 to: AAY21727 from: 1 to: 300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Amino acid sequence of rev-caspase-6.
                                                                                                                                                                                                                                                                                                     AAY21727 standard; Protein; 300 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 22; Fig 21C; 74pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (UYJE-) UNIV JEFFERSON THOMAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               cancer or autoimmune disease
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99WO-US00632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98US-0070987
                                                                                                                                                                                                                                                                                                                                                                                                                                 10-SEP-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-674-779-4/rev x AAY21727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            88.235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          99.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-419353/35.
                                                                                                                                                                                                                                                                         seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           300 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAX81216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              W09935277-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11-JAN-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               09-JAN-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                15-JUL-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Alnemri ES;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           alignment_scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             alignment_block:
                                                                                                                                                                                                                                                                                                                                                                      AAY21727;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seguence
```

```
60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                     The protein sequence was deduced from the DNA sequence of the expression vector pmTNF-MPH which was used for the expression of a fusion protein comprising the first 25 amino acids of murine TNF, a synthetic 6 His linker allowing for purification by metal affinity chromatography and a fragment of the 34 kD p362 protein from Mycobacterium paratuberculosis. The purified fusion protein may be used to assay for M. paratuberculosis infections, e.g. Johne's disease in cattle or other animals and possibly crohn's disease in humans. See also AAR27491-3.
                                                                                                   seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1992.DAT:AAR27492
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Polypeptide(s) derived from mycobacterium para-tuberculosis · are immunogens, useful as vaccines and for diagnosing Crohns disease
                                                                                                                                                                                                                                                                                                                                                                                         "synthetic 6 His site for metal affinity
                                                                                                                                                                                                                                                            Mycobacterium; Crohn's disease; Johne's disease; cattle; human; M. paratuberculosis; TNF.
    /note= "mouse TNF N-terminal fragment"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps: 1
Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gilot P;
                                                                                                                                                                                                                                   Fusion protein expressed by pmTNF-MPH.
                                                                                                                                                                                                                                                                                                                                                                                            /note= "synthetic 6 Hi
35..147
/note= "p326 protein"
                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                 AAR27492 standard; Protein; 147 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; Fig 9; 105pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          De KESELM,
                                                                                                                                                                                                                                                                                                        Mycobacterium paratuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    91EP-0400798
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       92WO-EP00661
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (INNO-) INNOGENETICS NV SA.
                                                                                                                                                                                                        05-MAR-1993 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4.300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Coene M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 1992-349213/42
                                                                                                                                    seg_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               147 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24-MAR-1992;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   25-MAR-1991;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09216628-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-OCT-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cocito C,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                           115 u 115
                                                                                                                                                                             AAR27492;
                                                 10 C 10
                                                                                                                                                                                                                                                                                                                                                    Peptide
                                                                                                                                                                                                                                                                                                                                                                                Peptide
                                                                                                                                                                                                                                                                                                                                                                                                          Protein
```

to: 147

to: AAR27492 from: 1

Align seg 1/1

US-09-674-779-4/rev x AAR27492

alignment\_block:

```
The present sequence represents a fusion protein comprising a Hepatitis C virus (HCV) NS3 protein. The NS3 polypeptides are used in a solid phase immunoassay comprising a HCV antigen (preferably a NS3 helicase or NS3 protease protein) in the presence of reducing agent on the solid phase. Use of the reducing agent and purification of the antigen using antibodation and desulphonation steps increases its reactivity to HCV antibodies and enables an earlier detection of HCV infection. The assay is used for detecting antibodies raised against the HCV antigen. The polypeptides are used for preventing and treating HCV infection. The polypeptides are also used for diagnosing hepatitis infection. The antibodies to these polypeptides are used for diagnosing hepatitis infection. The antibodies to these polypeptides are used for providing passive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   //note= "these residues represent the non-NS3
sequence, which is the mINF fusion partner,
the hexahistidine tag and part of the
                                                                                                                                                                                                   seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:AAY43894
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               New hepatitis C-virus polypeptide used for treating the infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HCV; NS3 helicase; HCV subtype la, HCV subtype lb, HCV infection; solid phase immunoassay; HCV antigen; NS3 protease; HCV antibody;
1 HisglnValGluGluGluGluGlyIleHisHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Zrein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Amino acid sequence of the mTNFH6NS3 fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   vaps: 1
Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sablon E,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                multilinker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Bosman A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Location/Qualifiers
                                                                                                                                                                                                                                                                                                 AAY43894 standard; Protein; 318 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Example 1; Fig 2-2; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99WO-EP02547.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      98EP-0870087
                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                4.300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Louwagie J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64.50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     passive vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     WPI; 2000-013283/01
                                                                                                                                                                                                                                                                        seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    318 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Hepatitis C virus.
                                                                                                                                                  36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                          10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAZ36163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9954735-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              15-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          L7-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                       11-FEB-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Maertens G,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      28-OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Synthetic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                            AAY43894;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Region
```

```
The present sequence represents a fusion protein comprising a Hepatitis C virus (HCV) NS3 protein. The NS3 polypeptides are used in a solid phase humunoassay comprising a HCV antigen (preferably a NS3 helicase or NS3 protease protein) in the presence of reducing agent on the solid phase. Use of the reducing agent and purification of the antigen using antibodies and enables an earlier detection of HCV infection. The assay is used for detecting antibodies raised against the HCV antigen. The polypeptides are used for preventing and treating HCV infection. The antibodies real enables and enables and enables and entibodies are used for preventing and treating HCV infection. The antibodies to these polypeptides are used for diagnosing hepatitis infection. The antibodies to these polypeptides are used for preventing and preating infection. The antibodies to these polypeptides are used for providing passive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /note= "these residues represent the non-NS3 sequence, which is the mINF fusion partner, the hexahistidine tag and part of the
                                                                                                                                                                                                                                  seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:AAY43895
                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV; NS3 helicase; HCV subtype la; HCV subtype lb; HCV infection; solid phase immunoassay; HCV antigen; NS3 protease; HCV antibody; passive vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  New hepatitis C-virus polypeptide used for treating the infection
                                                                                                                                                                                                                                                                                                                                                                                                 Amino acid sequence of the mTNFH6NS3 clone B9 fusion protein.
                                                                                       60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                            1 HisGlnValGluGluGluGlnGlyIleHisHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sablon E,
                                                     to: AAY43894 from: 1 to: 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            multilinker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bosman A,
                                                                                                                                                                                                                                                                      seq_documentation_block:
ID AAY43895 standard; Protein; 318 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Example 2; Fig 3-2; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     98EP-0870087.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99WO-EP02547.
alignment_block:
US-09-674-779-4/rev x AAY43894
                                                                                                                                                                                                                                                                                                                                                             11-FEB-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Maertens G, Louwagie J,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (INNO-) INNOGENETICS NV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 WPI; 2000-013283/01,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                   36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               318 AA;
                                                                                                                                                             10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-PSDB; AAZ36164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            W09954735-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     17-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              28-OCT-1999.
                                                       Align seg 1/1
                                                                                                                                                                                                                                                                                                                             AAY43895;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Region
```

Zrein M;

```
The present sequence represents a fusion protein comprising a Hepatitis C virus (HCV) NS3 protein. The NS3 polypeptides are used in a solid phase humunoassay comprising a HCV antigen (preferably a NS3 helicase or NS3 protease protein) in the presence of reducing agent on the solid phase. Use of the reducing agent and purification of the antigen using sulphonation and desulphonation steps increases its reactivity to HCV antibonation and antibodies and enables an earlier detection of HCV infection. The assay is used for detecting antibodies raised against the HCV antigen. The polypeptides are used for preventing and treating HCV infection. The polypeptides are also used for diagnosing hepatitis infection. The antibodies to these polypeptides are used for preventing and treating HCV infection. The antibodies to these polypeptides are used for proventing and preparation and passive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sequence, which is the mTNF fusion partner, the hexahistidine tag and part of the multilinker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      New hepatitis C-virus polypeptide used for treating the infection -
                                                                                                                                                                                                                                                                                                                    seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAY43898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HCV; NS3 helicase; HCV subtype la; HCV subtype lb; HCV infection; solid phase immunoassay; HCV antigen; NS3 protease; HCV antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Amino acid sequence of the mTNFH6NS3 Type 2a fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "these residues represent the non-NS3
                                                                                                                                                                                           21 HisGlnValGluGlnGlylleHisHisHisHisHisHis...valAs 36
                                                                                                                                                                        60 CATGTCTTACCAAAAATAAGCTACACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Zrein M;
                                                      Percent Identity: 55,000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sablon E,
                      Length:
                                         Gaps:
                                                                                                                                        Align seg 1/1 to: AAY43895 from: 1 to: 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bosman A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                      seq_documentation_block:
ID AAY43898 standard; Protein; 318 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Example 5; Fig 6-2; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             99WO-EP02547.
                                                                                  alignment_block:
US-09-674-779-4/rev x AAY43895
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               98EP-0870087
                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                     4.300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Maertens G, Louwagie J,
                  64.50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (INNO-) INNOGENETICS NV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        passive vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2000-013283/01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hepatitis C virus.
                                                                                                                                                                                                                                                                                 36 pProGlyPro 39
                                                                                                                                                                                                                                                 10 CCCTCTGCCT 1
                                   Ratio:
Percent Similarity:
                  Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; AAZ36167
                                                                                                                                                                                                                                                                                                                                                                                                                                               11-FEB-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  W09954735-A1
alignment_scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             .7-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28-OCT-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Synthetic.
                                                                                                                                                                                                                                                                                                                                                                                                            AAY43898;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Region
```

SXCC

```
Quality:
Ratio:
                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO9954735-A1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Maertens G,
                                                                     vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11-FEB-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28-0CT-1999
                                                                                                                                             alignment_scores
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAY43900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8 X S S S S S X 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The present sequence represents a fusion protein comprising a Hepatitis C virus (HCV) NS3 protein. The NS3 polypeptides are used in a solid phase immunoassay comprising a HCV antigen (preferably a NS3 helicase or NS3 protease protein) in the presence of reducing agent on the solid phase. Use of the reducing agent and purification of the antigen using sulphonation and desulphonation steps increases its reactivity to HCV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note= "these residues represent the non-NS3 sequence, which is the mINF fusion partner, the hexahistidine tag and part of the
                                                                                                                                                                                                                                                                                                                                   seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:AAY43899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              New hepatitis C-virus polypeptide used for treating the infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           HCV; NS3 helicase; HCV subtype la: HCV subtype lb; HCV infection; solid phase immunoassay; HCV antigen; NS3 protease; HCV antibody; passive vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Amino acid sequence of the mTNFH6NS3 Type 2b fusion protein.
                                                                                                                                                                                                                    60 CATGICITACCAAAAATAAGCIACACCATCACCAICACCAITAAGICGA 11
                                                                                                                                                                                                                                    1 HisGlnValGluGluGluGlyIleHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Zrein
                                                                                                                Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Bosman A, Sablon E,
                                                                                      Length:
                                                                                                       Gaps:
                                                                                                                                                                                          to: 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               multilinker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                  AAY43899 standard; Protein; 318 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 6; Fig 7-2; 66pp; English.
                                                                                                                                                                                          Align seg 1/1 to: AAY43898 from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      99WO-EP02547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 98EP-0870087
                                                                                                                                                 alignment_block:
US-09-674-779-4/rev x AAY43898
                                                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Louwagie J,
                                                                                                     4.300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (INNO-) INNOGENETICS NV
                                                                                       64.50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-013283/01.
N-PSDB; AAZ36168.
                                                                                                                                                                                                                                                                                                                                                                    seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Synthetic.
Hepatitis C virus.
                                318 AA;
                                                                                                                                                                                                                                                                                                            36 pProGlyPro 39
                                                                                                                                                                                                                                                                               10 CCCTCTGCCT 1
                                                                                                                   Percent Similarity:
                                                                        alignment_scores:
Quality:
                                                                                                       Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    15-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            W09954735-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                            11-FEB-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28-OCT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Maertens G,
   vaccination.
                                Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                 AAY43899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Key
Region
```

```
The present sequence represents a fusion protein comprising a Hepatitis C
antibodies and enables an earlier detection of HCV infection. The assay is used for detecting antibodies raised against the HCV antigen. The polypeptides are used for preventing and treating HCV infection. The polypeptides are also used for diagnosing hepatitis infection. The antibodies to these polypeptides are used for providing passive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "these residues represent the non-NS3 sequence, which is the mINF fusion partner, the hexahistidine tag and part of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:AAY43900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   New hepatitis C-virus polypeptide used for treating the infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HCV: NS3 helicase; HCV subtype la; HCV subtype lb; HCV infection; solid phase immunoassay; HCV antigen; NS3 protease; HCV antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Amino acid sequence of the mTNFH6NS3 Type 2c fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 HisglnValgluGluGlnGlyIleHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Zrein M;
                                                                                                                                                                                                                                                                                                                                                                  Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ы
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Bosman A, Sablon
                                                                                                                                                                                                                                                                                                                                              Gaps:
                                                                                                                                                                                                                                                                                                                 Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          to: 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          multilinker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          seq_documentation_block:
ID AAY43900 standard; Protein; 318 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Example 7; Fig 8-2; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Align seg 1/1 to: AAY43899 from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99WO-EP02547.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98EP-0870087
                                                                                                                                                                                                                                                                                                                                                                                                                            alignment_block:
US-09-674-779-4/rev x AAY43899
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Louwagie J,
                                                                                                                                                                                                                                                                                                                                     4.300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (INNO-) INNOGENETICS NV
                                                                                                                                                                                                                                                                                                                    64.50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI: 2000-013283/01.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                passive vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                       Sequence 318 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     N-PSDB; AAZ36169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10 CCCTCTGCCT 1
```

```
virus (HCV) NS3 protein. The NS3 polypeptides are used in a solid phase immunoassay comprising a HCV antigen (preferably a NS3 helicase or NS3 protease protein) in the presence of reducing agent on the solid phase. Sulphonation and desulphonation of the antigen using antibodies and enables an earlier detection of HCV infection. The assay la used for detecting antibodies raised against the HCV antigen. The polypeptides are used for preventing and treating HCV infection. The antibodies are also used for diagnosing hepatitis infection. The antibodies are also used for diagnosing hepatitis infection. The vaccination and the end of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Amino acid sequence of the mTNFH6NS3 Type 3a clone 21 fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             sequence, which is the mINF fusion partner,
the hexahistidine tag and part of the
multilinker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAY43896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HCV; NS3 helicase; HCV subtype la; HCV subtype lb; HCV infection; solid phase immunoassay; HCV antigen; NS3 protease; HCV antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note= "these residues represent the non-NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1 HisGlnValGluGluGluGlnGlyIleHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Zrein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Louwagie J, Bosman A, Sablon E,
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          to: 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            seq_documentation_block:
ID AAY43896 standard; Protein; 326 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Align seg 1/1 to: AAY43900 from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               alignment_block:
US-09-674-779-4/rev x AAY43900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       99WO-EP02547.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98EP-0870087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         11-FEB-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           4.300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (INNO-) INNOGENETICS NV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                passive vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-013283/01.
N-PSDB; AAZ36165.
                                                                                                                                                                                                                                                                                                                                   318 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                               Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                               alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9954735-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       15-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       17-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-OCT-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Maertens G,
                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAY43896;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Key
Region
     22222222222222
```

```
The present sequence represents a fusion protein comprising a Hepatitis (virus (HCV) NS3 protein. The NS3 polypeptides are used in a solid phase bimunoasasy comprising a HCV antigen (preferably a NS3 helicase or NS3 protease protein) in the presence of reducing agent on the solid phase. Sulphonation and desulphonation steps increases its reactivity to HCV antibonation and desulphonation steps increases its reactivity to HCV antibonation and enables an earlier detection of HCV infection. The assay is used for detecting antibodies raised against the HCV antigen. The polypeptides are used for preventing and treating HCV infection. The antibodies to these polypeptides are used for diagnosing hepatitis infection. The antibodies to these polypeptides are used for greventing and treating HCV infection. The antibodies to these polypeptides are used for grandless for providing passive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Amino acid sequence of the mTNFH6NS3 Type 3a clone 32 fusion protein.
New hepatitis C-virus polypeptide used for treating the infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sequence, which is the mTNF fusion partner, the hexahistidine tag and part of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:AAY43897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HCV; NS3 helicase; HCV subtype la; HCV subtype lb; HCV infection; solid phase immunoassay; HCV antigen; NS3 protease; HCV antibody;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          /note= "these residues represent the non-NS3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   21 HisGlnValGluGluGluGluGlnGlyIleHisHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length: 20
Gaps: 1
Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              to: AAY43896 from: 1 to: 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          multilinker"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        seq_documentation_block:
ID AAY43897 standard; Protein; 326 AA.
                                                                                                             Example 4; Fig 4-2; 66pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 99WO-EP02547.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                alignment_block:
US-09-674-779-4/rev x AAY43896
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    98EP-0870087
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11-FEB-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       passive vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      326 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Synthetic.
Hepatitis C virus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              vaccination.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09954735-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    17-APR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28-OCT-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Align seg 1/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AAY43897;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Key
Region
      \begin{smallmatrix} \mathbf{F} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ \mathbf{X} & \mathbf{X} & \mathbf{X} \\ \mathbf{X} & \mathbf{X}
```

```
27-JAN-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The present sequence represents a fusion protein comprising a Hepatitis C virus (HCV) NS3 protein. The NS3 polypeptides are used in a solid phase finanuosassay comprising a HCV antigen (preferably a NS3 helicase or NS3 protease protein) in the presence of reducing agent on the solid phase. Use of the reducing agent and purification of the antigen using
                                                                                                                                                                                                                                                                                                                                        sulphonation and desulphonation steps increases its reactivity to HCV antibodies and enables an earlier detection of HCV infection. The assay is used for detecting antibodies raised against the HCV antipon. The polypeptides are used for preventing and treating HCV infection. The polypeptides are also used for diagnosing hepatitis infection. The antibodies to these polypeptides are used for many expectation and the providing passive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR97229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Brucellosis; vaccine; antigen; diagnosis; ELISA; Escherichia coli.
                                                                                                                      New hepatitis C-virus polypeptide used for treating the infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
     Zrein M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps: 1
Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1..37
/note= "vector-derived sequence"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "Brucella 17 kDa antigen"
  Sablon E,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Align seg 1/1 to: AAY43897 from: 1 to: 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Recombinant Brucella 17 kDa antigen.
     Maertens G, Louwagie J, Bosman A,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 seq_documentation_block:
ID AAR97229 standard; Protein; 195 AA.
                                                                                                                                                                             Example 4; Fig 5-2; 66pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               95WO-EP04670.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94EP-0870185.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-674-779-4/rev x AAY43897
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Chimeric Brucella abortus;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            64.50
4.300
75.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    38..195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Chimeric synthetic.
                                                      WPI; 2000-013283/01.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  326 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     10 CCCTCTGCCT 1
                                                                                N-PSDB; AAZ36166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25-NOV-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO9617065-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               27-NOV-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20-0CT-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 06-JUN-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 alignment block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAR97229;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME OF COLORS O
```

```
A fusion protein (AAR97229) is composed of an N-terminal 37-amino acid sequence derived from vector pIGFIO (including a hexahistidine leader peptide), and the 17 kba antigen (see also AAR97225) of Brucella abortus. It was obtd. following PCR amplification (see also AA728817-18) of the 17 kba antigen gene (AA728816), its insertion into pIGFHIO and expression in Sechetichia coll. The fusion protein was used in competition ELISAs to study the reaction of Brucella positive sera. It can also be used in indirect ELISA to determine anti-Brucella antibodies present in a sample.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:AAW26698
                                                                                                                                                       Isolated 17kD Brucella antigen - used to develop prods. to detect, and provide protective immunity against Brucella infection
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tg20; antigen; toxoplasmosis; infection; therapy; diagnosis;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 HisGlnValGluGlnGlyIleHisHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          60 CATGICITACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length: 19
Gaps: 1
Percent Identity: 52.632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /note= "mouse INF leader peptide"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Toxoplasma gondii antigen Tg20 fusion protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29..37
/note= "hexahistidine tag"
38..252
/note= "Tg20"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Align seg 1/1 to: AAR97229 from: 1 to: 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             vaccine; tumour necrosis factor; TNF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        seq_documentation_block:
ID AAW26698 standard; Protein; 272 AA
                                                                                                                                                                                                                                        Example 2; Fig 3; 69pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Chimeric - synthetic.
Chimeric - Toxoplasma gondii.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              97WO-EP00394.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              alignment_block:
US-09-674-779-4/rev x AAR97229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27-MAR-1998 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4.233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Chimeric - Mus musculus.
(INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    63.50
                                                                                                          WPI; 1996-277783/28
                                                          Saman E;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 195 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36 pProMet 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10 CCCTCTG 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WO9727300-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31-JUL-1997
                                                          Hemmen F,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AAW26698;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Protein
```

```
New isolated Toxoplasma gondii antigen Tg20 - used to develop products for the diagnosis and prevention of {\bf T}_{\cdot} gondii infection in
                                                                                                                                                                                                                                                                                                                           This sequence comprises a fusion protein composed of the mouse tumour necrosis factor leader peptide, a polyhistidine tag, and antigen Tg20 (see AAW26692) of Toxoplasma gondii. A vector containing DNA encoding the fusion protein was used to transform Escherichia coli MNIO61 host cells for production of recombinant Tg20. Antigen Tg20 polypeptides and immunodominant peptides (see AAW26693-97) can be used in claimed methods for the diagnosis and prevention of T. gondii infection in mammals.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AAG67271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Amino acid sequence of the S.aureus NAD synthetase (NadE).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60 CATGICTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGICGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 HisGlnValGluGluGlnGlylleHisHisHisHisHis...valAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Identity: 55.556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAD synthetase; NadE; protein co-ordinate data.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      to: AAW26698 from: 1 to: 272
                                                                                                              Van Heuverswyn H;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               seq_documentation_block:
ID AAG67271 standard; protein; 282 AA.
                                                                                                                                                                                                                                                                                   Example 3; Fig 4; 68pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (PHAA ) PHARMACIA & UPJOHN CO.
                       96EP-0870006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  31-JAN-2000; 2000US-0179261.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               alignment_block:
US-09-674-779-4/rev x AAW26698
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        30-JAN-2001; 2001WO-US02913
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                13-NOV-2001 (first entry)
                                                               (INNO-) INNOGENETICS NV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61.50
4.393
77.778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Prince DB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Staphylococcus aureus
                                                                                                            Saman E,
                                                                                                                                                    WPI; 1997-393688/36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-488798/53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      272 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO200155443-A1.
                     26-JAN-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             36 pPro 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              02-AUG-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10 CCCT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Benson TE,
                                                                                                            Jacobs D,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Align seg 1/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence
                                                                                                                                                                                                                                                  mammals
A A C A A C A A C A A C A A C A A C A A C A A C A A C A A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A C A
```

```
A molecule comprising an S.aureus NAD synthetase or NAD synthetase-like substrate binding pocket with defined structure coordinates for identifying inhibitors of NAD synthetase for rational drug design -
                                                                                                                                                                            The present sequence represents a NAD synthetase, designated NadE, which is isolated from Staphylococcus aureus. The NadE polypeptide is defined by a set of points with a root mean square deviation of less than about 1.1 Angstrom from points representing the backbone of the amino acids as represented by structure coordinates fully defined in the specification. The protein-coordinate data for NadE is useful for identifying and making an inhibitor of NadE activity. It is also useful for for drug discovery using rational drug design. The data is useful for solving the structure of other crystal forms of S.aureus NadE or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:AA011238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukaemia, inflammation and immune
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Identity: 72.727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        to: AAG67271 from: 1 to: 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               272 ProLysSerArgSerHisHisHisHis 282
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              51 CCAAAAATAAGCTACACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         AA011238 standard; Protein; 109 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human polypeptide SEQ ID NO 25130.
                                                                                                                          Claim 42; Fig 9; 940pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Tang YT, Liu C, Drmanac RT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28-FEB-2000; 2000US-0515126.
18-MAY-2000; 2000US-0577409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26-FEB-2001; 2001WO-US04927
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-674-779-4/rev x AAG67271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             06-NOV-2001 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S. aureus NadE complexes.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6.100
90.909
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       61.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2001-514838/56.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         282 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (HYSE-) HYSEQ INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; AAI91169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO200164835-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               07-SEP-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 disorders -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           alignment_block
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Align seg 1/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AA011238;
```

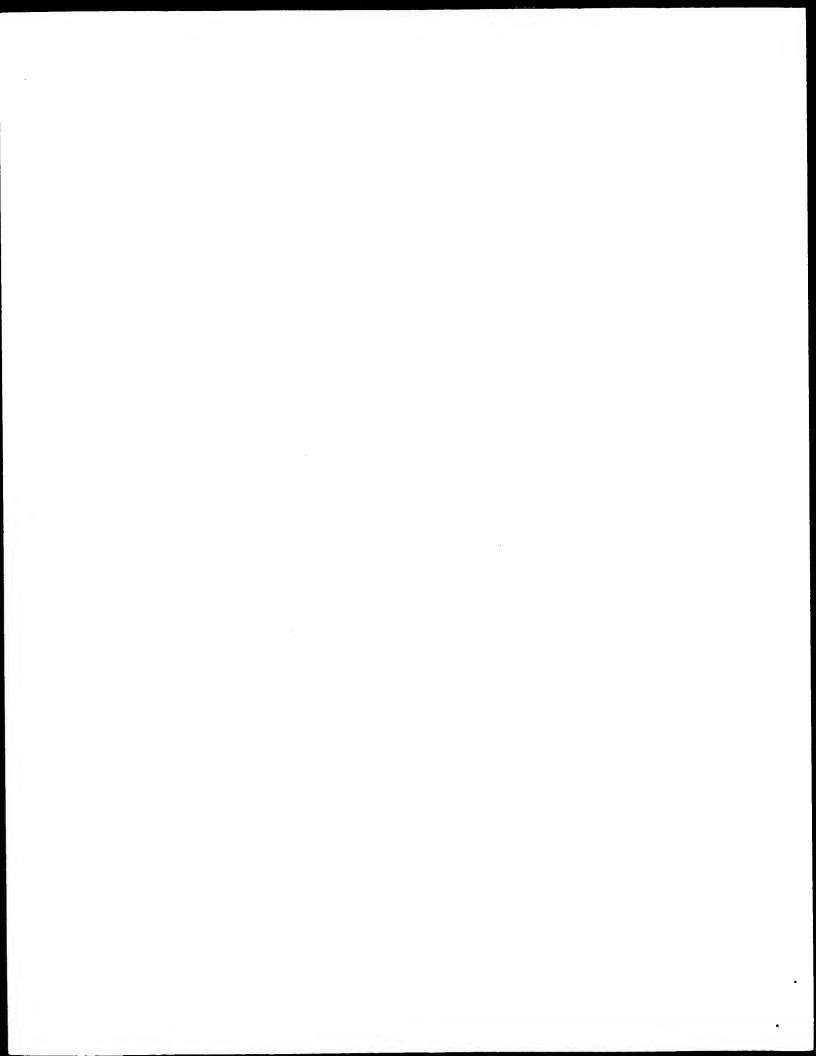
œ

```
Domain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO00010-AAO19910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, haematopoiesis regulating activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Interleukin; IL-17C.his; PRO1122 polypeptide; clone DNA62377-1381-1; immunoprecipitation; IL-17 receptor extracellular domain; IL-17R ECD; cytokine IL-17; hybridisation probe; antagonist; Gly(His)8 tag; agonist; degenerative cartilaginous disorder; diagnose; therapy.
                                                                                                                                                                                                                                 Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA2000.DAT:AAY44485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /label- Mature_IL-17C_polypeptide
/note= "Used to treat degenerative cartilaginous
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Human Interleukin 17C with C-terminal Gly(His)\theta tag, IL-17C.his.
 Claim 20; SEQ ID NO 25130; 1399pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Identity: 57.143
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note= "C-terminal Gly(His)8 tag"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /note= "Conserved Cys residue"
189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /note= "Conserved Cys residue"
Misc-difference 198..206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /note= "Conserved Cys residue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /note= "Conserved Cys residue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           'note= "Conserved Cys residue"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     note= "Conserved Trp residue"
                                                                                                                                                                                                                                                                         ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                 Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Align seg 1/1 to: AA011238 from: 1 to: 109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ..18
|abel= Signal_peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY44485 standard; Protein; 206 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           alignment_block:
US-09-674-779-4/rev x AAO11238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    disorder"
                                                                                                                                                                                                                                                                                                                                                                                                  5.455
                                                                                                                                                                                                                                                                                                                                                                                    60.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference 129
                                                                                                                                                                                                                                                                                                              Sequence 109 AA;
                                                                                                                                                                                                                                                                                                                                                                                    Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             27-MAR-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                     alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAY44485;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Protein
```

```
The present sequence is the human PRO1122 polypeptide, with a C-terminal Gly(His)8 tag, IL-17C.his, derived from the clone DNA62377-1381-1. This sequence is used in a competitive binding experiment for the immunoprecipitation of IL-17 receptor extracellular domain (IL-17R ECD). The entire coding region of IL-17C can be used as hybridisation probe. The PRO1122 polypeptide, agonist or antagonist, is used to diagnose and treat a degenerative cartilaginous disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1996.DAT:AAR95055
                                                                                                                                                                                                                                                                                                                                                           New polypeptides designated PRO1031 and PRO1122 used to treat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Nucleic acid transfer system; gene transfer; gene therapy; cell targeting; multidomain protein; vector; cancer; exotoxin A; DETA; ompA; signal peptide; GAL4; interleukin-2;
                                                                                                                                                                                                                                                                     Li H,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length: 13
Gaps: 0
Percent Identity: 69.231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     57 GTCTTACCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                     Goddard A, Gurney AL,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Align seg 1/1 to: AAY44485 from: 1 to: 206
                                                                                                                                                                                                                                                                                                                                                                                                                                            Example 11; Page 138-139; 141pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1..8
/label= FLAG_epitope
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IL-2-DETA-DGAL4 multidomain protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                             degenerative cartilaginous disorder
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              seq_documentation_block:
ID AAR95055 standard; Protein; 421 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chimeric synthetic;
Chimeric Homo sapiens;
Chimeric Pseudomonas aeruginosa;
Chimeric Saccharomyces cerevisiae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /label= Spacer
18..150
                                                                                              99WO-US10733.
                                                                                                                                          98US-0085579
                                                                                                                                                                       98US-0113621
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-674-779-4/rev x AAY44485
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           19-AUG-1996 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5.455
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          60.00
                                                                                                                                                                                                                                                                     Chen J, Filvaroff E,
                                                                                                                                                                                                                      (GETH ) GENENTECH INC
                                                                                                                                                                                                                                                                                                                      WPI; 2000-116314/10.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             alignment_scores:
WO9960127-A2
                                                                                              14-MAY-1999;
                                                                                                                                                                       23-DEC-1998;
                                               25-NOV-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAR95055;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          IL-2.
```

```
A multidomain protein (AAR95055) has a FLAG epitope, a portion of human interleukin-2 that acts as a ligand domain, a non-vytotoxic portion of Pseudomonas aeruginosa exotoxin A acting as a translocation domain and the DNA binding domain of yeast GAL4. It is the product of a fusion gene (AAT29411) and can be expressed in E. coli (resulting in removal of an ompA signal peptide). It is used with an effector nucleic acid that comprises e.g. a gene to be delivered to a cell and a cognate structure for the GAL4 DNA binding suitable for gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                    Nucleic acid transfer system for gene therapy, e.g. against cancer - includes toxin translocation domain to target nucleic acid to specific cell
                                                                                                                                               414..421
/label= Spacer
/note= "endoplasmic reticulum retention signal"
 /label- IL-2
/note= "amino acids 1-113 of human IL-2"
                                                                                                                        /label= GAL4
/note= "amino acids 2-147 of yeast GAL4"
                                                            /label= ETA
/note= "amino acids 252-366 of ETA"
                                                                                                                                                                                                                                                                                                                                                                                                                                 Claim 7; Page 67-69; 106pp; English.
                                    /label= Spacer
152..266
                                                                                                 'label= Spacer
                                                                                                                                                                                                                                                  95WO-EP04270.
                                                                                                                                                                                                                                                                       94EP-0810627.
                                                                                                                .413
                                                                                                                                                                                                                                                                                                                         Fominaya J, Wels W;
                                                                                                                                                                                                                                                                                                                                               WPI; 1996-239505/24.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  421 AA;
                                                                                                                                                                                                                                                                                                                                                             N-PSDB; AAT29411.
                                                                                                                                                                                                                                                                                                 (WELS/) WELS W.
                                                                                                                                                                                                                                                                       01-NOV-1994;
                                                                                                                                                                                               WO9613599-A1
                                                                                                                                                                                                                                                 31-OCT-1995;
                                                                                                                                                                                                                       09-MAY-1996.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                          Peptide
                                                                                    Peptide
                                                                                                                                                Peptide
                                                 Domain
                                                                                                             Domain
```

alignment\_scores:



```
Sequence 43, Application US/08447430A; Patent No. 591658
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 10, Application US/08244951A Patent No. 5843779 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      BIERMAN & MUSERLIAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION: 424
INFORMATION FOR SEQ ID NO: 43:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  600 THIRD AVENUE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        338 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                     TITLE OF INVENTION: Reco
TITLE OF INVENTION: DOLY
TITLE OF INVENTION: LUB
NUMBER OF SEQUENCES: 43
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    : protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64.50
4.300
75.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             amino acid
GY: linear
                                                                                                                                                                                                                                                                 seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      seq_documentation block
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CITY: NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HYPOTHETICAL:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-447-430A-43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1182
2265
2509
                                                                                                                                                                                                                                -WODEL-frame-holp model -DEV=xlp
-Q=Cgg02_1VGSPTO_spool/US09674779/runat_30072002_151752_7276/app_query.fasta_1.116
-Q=Cgg02_1VGSPTO_spool/US09674779/runat_30072002_151752_7276/app_query.fasta_1.116
-DB=ISSued_Patents_AA -OFM#-fastan -SUFFIX=xai -GADPD=12.000
-GAPEXT=4.000 -MINMATCH=0.00 -LCOPCL=0.000 -LCOPEXT=0.000
-GGAPOP=4.500 -GGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-GGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -XGAPEXT=0.500
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-TRANS=human40.cdi -LIST=49 -DOCALIGN=200 -THR_SCORE=pct
-TRANS=human40.cdi -LIST=49 -NOCALIGN=200 -THR_SCORE=pct
-TOCALIGN=120 -MINLEN=0 -MAXLEN=2000000000
-USER=US09674779_GCGN1_1_29 -NCPU=6 -LCPU=3 -LONGLOG
-DEV_TIMBGOUT=120 -WARN_TIMEOUT=30 -NO_XLPXX -WAIT -THREADS=1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence

/cgn2_6/prodata/2/laa/SB_COMB. pep:105-08-417-430A-13
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-08-417-430A-13
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-08-417-430A-13
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-08-417-430A-13
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-08-418-917-23
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-09-418-952A-23
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-09-418-952A-23
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-09-418-952A-23
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-09-1011-135-0
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-09-1011-135-0
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-09-1011-135-0
/cgn2_6/prodata/2/laa/SB_COMB. pep:105-09-275-2
/cgn2_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 0.5969
                                                                                                                    About: Results were produced by the GenCore software, version 4.5, Copyright (c) 1993-2000 Compugen Ltd.
                  out_format : pfs
OM of: US-09-674-779-4 to: Issued_Patents_AA:*
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Database: Issued_Patents_AA:*
Database sequences: 211628
Database length: 24425594
Search time (sec): 22.710000
                                                             Date: Jul 30, 2002 4:25 PM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Search information block:
Query: US-09-674-779-4
                                                                                                                                                                                                          Command line parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query length: 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                score_list:
```

```
0.8453
0.8453
0.8942
0.9453
                                                                                                                                                                                                                                                                                                                                                                                            Recombinant polypeptides and peptides, nucleic acids coding for the same and use of these polypeptides and peptides in the diagnostic of tuberculosis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;
APPLICANT: VANMECHELEN, EUGEEN; VAN DE VOORDE, ANDRE
TITLE OF INVENTION: MONOCLONAL ANTIBODIES
TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
TITLE OF INVENTION: PROTEIN TAU, HYBRIDOMAS SERRETING THESE
TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATIONS
            163.08 (163.08 (158.85) 158.85
         52.00
52.00
52.00
52.00
52.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  60 CATGTCTTACCAAAAAAAAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21 HisGlnValCluGluGlnGlyIleHisHisHisHisHis...valAs 36
                                                                                                                                                                                 seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-447-430A-43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-244-951A-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO) CURRENT APPLICATION DATA: APPLICATION NUMBER: US/08/447,430A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Align seg 1/1 to: US-08-447-430A-43 from: 1 to: 338
/cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-226-264-20 - /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-226-264-8 - /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-226-264-21 - /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-895-707-21 - /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-664-449-39 - /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-664-449-39 -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      alignment_block:
US-09-674-779-4/rev x US-08-447-430A-43
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
```

COUNTRY:

```
DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED PROTEIN TAD, HYBRIDDARS SECRETING THESE ANTIBODIES, ANTIGEN RECOGNITION BY THESE MONOCLONAL ANTIBODIES AND THEIR APPLICATIONS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21 HisGlnValGluGlnGlyIleHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              60 CATGTCTTACCAAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-389-011-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: VANDERMEBREN, MARC; MERCKEN, MARC; ANDRE TITLE OF INVENTION: MONOCLONAL ANTIBODIES TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-TITLE OF INVENTION: PROFEIN TAD, HYBRIDOMAS SECRETING TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY TITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 to: US-08-244-951A-10 from: 1 to: 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: mTHFMPH-taul fusion protein
                                                                                                                                                             SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/244,951A
FILING DATE: 19-JAN-1995
CLASSIFICATION: 435
PRIOR APPLICATION 9435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/03499
FILING DATE: 10-DEC-1993
PRIOR APPLICATION DATE: BP/92/403403.6
FILING DATE: 14-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: CHARLES A: MUSERLIAN
REGISTRATION NUMBER: 19,683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REGISTRATION NUMBER: 19,683
REFERENCE/DOCKET NUMBER: 410.003A
TELECOMONICATION INFORMATION:
TELEPHONE: (212) 661-8000
TELEFAX: (212) 661-8002
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  alignment_block:
US-09-674-779-4/rev x US-08-244-951A-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5: BIERMAN & MUSERLIAN 600 THIRD AVENUE
                                                                                                                COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 23, Application US/08389011 Patent No. 5861257 GENERAL INFORMATION:
                                                                                            FLOPPY DISK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                64.50
4.300
75.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES: 2 CORRESPONDENCE ADDRESS:
                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Amino Acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seg_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                             10016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-08-244-951A-10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Align seg 1/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET:
                         COUNTRY:
```

```
APPLICANT: VANDERMEEREN, MARC; MERCKEN, MARC;
APPLICANT: VANDECHELBN, EUGEBN;
APPLICANT: VAN DE VOORDE, ANDER
TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
TITLE OF INVENTION: DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED
TITLE OF INVENTION: PROPIEIN TAU, HYBRIDOMAS SECRETING THESE
TITLE OF INVENTION: ANTIBODIES, ANTIGEN RECOGNITION BY THESE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            seq_name: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-08-403-917A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Align seg 1/1 to: US-08-389-011-23 from: 1 to: 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: CHARLES A. MUSERLIAN
REGISTRATION NUMBER: 19,683
REFERENCE/DOCKET NUMBER: 410.003-1-CON
TELECOMUNICATION INFORMATION:
TELEFAX: (212) 661-8000
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                 APPLICATION NUMBER: US/08/389,011
FILING DATE: 15-FEB-1995
CLASSIPFCATION: 435
PROOR APPLICATION DATA:
APPLICATION NUMBER: 08/403,917
FILING DATE: 19-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/403,916
FILING DATE: 19-JAN-1995
                                                                                                                                                                                                                                                                                                                                                             PILING APPLICATION DATA:
APPLICATION NUMBER: 08/244,951
FILING DATE: 13-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/03499
FILING DATE: 10-DEC-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: EP/92/403403.6 FILING DATE: 14-DEC-1992 ATTORNEY/AGENT_INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-674-779-4/rev x US-08-389-011-23
                                                                                       PC-DOS/MS-DOS
              COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4.300
                                                                                                                              CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; TOPOLOGY: Unknown
US-08-389-011-23
                                                                                  OPERATING SYSTEM:
SOFTWARE: ASCII
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: Amino Acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Quality:
10016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     alignment_block:
```

```
IITLE OF INVENTION: MONOCLONAL ANTIBODIES AND THEIR APPLICATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  21 HisGlnValGluGluGlnGlyIleHisHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   seq_name: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-348-952A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Align seg 1/1 to: US-08-403-917A-23 from: 1 to: 391
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: VANDERMEBREN, MARC; MERCKEN, MARC; APPLICANT: VANMECHELEN, EUGEEN; APPLICANT: VAN DE VOORDE, ANDRE TITLE OF INVENTION: MONOCLONAL ANTIBODIES
                                                                                                                                                                                                                                                                         SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/403,917A
FILING DATE: 19-JAN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/256,167
FILING DATE: 27-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/244,951
FILING DATE: 13-JUN-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER: EP/92/403403.6 FILING DATE: 14-DEC-1992 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP93/03499
FILING DATE: 10-DEC-1993
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME: CHARLES A. MUSERLIAN
RECISTRATION NUMBER: 19,683
REFERENCE/DOCKET NUMBER: 410.003-1
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                alignment_block:
US-09-674-779-4/rev x US-08-403-917A-23
                                                                                                                                                                                                               MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
                      NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIERMAN & MUSERLIAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Sequence 23, Application US/09348952A Patent No. 6232437 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEPHONE: (212) 661-8000
TELEFAX: (212) 661-8002
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
                                                   ADDRESSEE: BIEKMAN, -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    64.50
4.300
75.000
                                                                                                                                                                                       COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Amino Acid
                                                                                                        NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    36 pProGlyPro 39
                                                                                                                                                   USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                             MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ratio:
                                                                                                                                                 COUNTRY: US ZIP: 10016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; TOPOLOGY:
US-08-403-917A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              alignment_scores:
                                                                                                                             STATE:
```

```
DIRECTED AGAINST THE MICROTUBULE-ASSOCIATED PROTEIN THY, HYBRIDDMAS SECRETIC THESE ANTHERDES, ANTHER BY THESE MONOCLONAL ANTIGODIES AND THEIR APPLICATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            60 CATGICITACCAAAAAAAAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 HisGlnValGluGluGluGlnGlyIleHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq_name: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-849-634B-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 64.50 Length: 20
4.300 Gaps: 1
75.000 Percent Identity: 55.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           to: US-09-348-952A-23 from: 1 to: 391
                                                                                                                                                                                                                                                                                                                                                                                                PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/403,917
FILING DATE: 19-JAN-1995
APPLICATION NUMBER: 08/256,167
FILING DATE: 27-JUN-1994
FILING DATE: 27-JUN-1994
PRIOR APPLICATION DATA: 3-JUN-1994
FILING DATE: 13-JUN-1994
PRIOR APPLICATION NUMBER: 08/244,951
FILING DATE: 10-DEC-1993
PRIOR APPLICATION NUMBER: EP/92/403403.6
FILING DATE: 14-DEC-1993
PRIOR DATE: 14-DEC-1992
ATONDEY-AGENT INFORMATION:
NAME: CHARLES A. MUSERLIAN
REGISTRATION NUMBER: 19,683
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: 19,683
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 661-8000
                                                                                                                                                                                                                                                                                                                                         CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/348,952A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           alignment_block:
US-09-674-779-4/rev x US-09-348-952A-23
                                                                                                                                                                                                                                                                                                         PC-DOS/MS-DOS
                                                                                                                      SEE: BIERMAN & MUSERLIAN
: 600 THIRD AVENUE
NEW YORK
NEW YORK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; Sequence 3, Application US/08849634B; Patent No. 6296855
                                                                                                                                                                                                                                                               E: FLOPPY DISK
IBM PC COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELEFAX: (212) 661-8002
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
                                                                           NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIERMAN &
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Unknown
                                                                                                                                                                                                                                        COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                    COMPUTER: IBM PC (OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY: Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Amino Acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       seq_documentation_block:
                                                                                                                                                                                                                                                                                                                            ASCII
                                                                                                                                                                                                      USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36 pProGlyPro 39
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10 CCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:
                                                                                                                                                                                                                     10016
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-348-952A-23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              alignment_scores:
                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                          SOFTWARE:
                                                                                                                                         STREET:
CITY: NE
STATE: N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Align seg 1/1
```

us-09-674-779-4.rai

```
Taegu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Seoul
                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |||||
36 pPro 37
                                                                                                                                                                                                                                   US-09-101-135-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10 CCCT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CITY: STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STATE:
  APPLICANT: INNOGENETICS N.V.

TITLE OF INVENTION: NEW 17 kDa BRUCELLA ABORTUS ANTIGEN, RECOMBINANT
TITLE OF INVENTION: NEW 17 kDa BRUCELLA ABORTUS ANTIGEN, RECOMBINANT
TITLE OF INVENTION: THEREOF IN DIAGNOSTIC AND PROPHYLACTIC METHODS AND KITS
FILE REPERBNCE: INNSOO6
CURRENT APPLICATION NUMBER: US/08/849,634B
CURRENT FILING DATE: 1997-12-01
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PATENTIN VOF: 2.1
SEQ ID NO 3: 8
                                                                                                                                                                                                                                                                                              OTHER INFORMATION: Description of Artificial Sequence:SYNTHETIC OTHER INFORMATION: CONSTRUCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 HisGlnValGluGluGluGlyIleHisHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    60 CATGTCTTACCAAAAAATAAGCTACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              seq_name: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-09-101-135-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: JACOBS, DIRK
APPLICANT: SAMAN, ERIC
APPLICANT: SAMAN, ERIC
TITLE OF INVENTION: TOXOPLASMA GONDII ANTIGEN T920
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Identity: 52.632
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SOFTWARE: Patentin Release #1.0, Version #1.30 CURRENT APPLICATION DATA: APPLICATION NUMBER: US/09/101,135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Align seg 1/1 to: US-08-849-634B-3 from: 1 to: 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: NIXON & VANDERHYE P.C. STREET: 1100 NORTH GLEBE ROAD, 8TH FLOOR
                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER: EP 96870006.2 FILING DATE: 26-JAN-1996 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO 97/27300
FILING DATE: 27-JAN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      alignment_block: US-09-674-779-4/rev \times US-08-849-634B-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                       TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                         4.233
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                      63.50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SADOFF, B.J.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ARLINGTON
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                    alignment_scores:
Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        36 pProMet 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10 CCCTCTG 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ٧
                                                                                                                                                                                                                                                                                                                                        US-08-849-634B-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                      LENGTH: 195
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STATE:
                                                                                                                                                                                                                                                                                     FEATURE:
```

```
GENERAL INFORMATION:
APPLICANT: SHIN, Hang Cheol
APPLICANT: CHANG, Seung Gu
APPLICANT: CHANG, Seung Gu
APPLICANT: KIM, Dae Young
APPLICANT: KIM, Chong Suhl
TITLE OF INVENTION: Proinsulin Derivative and Process
TITLE OF INVENTION: Proinsulin Derivative and Process
TITLE OF INVENTION: Proinsulin Derivative and Process
TITLE OF INVENTION: Proinsulin NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSE: SHIN, Hang Cheol
STREET: $200 MB ADDRESSES SHIN, Hang Cheol
STREET: $200 MB ADDRESSES SHIN, Hang Cheol
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 HisGlnValGluGluGlnGlyIleHisHisHisHisHis...ValAs 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       60 CATGICITACCAAAAAAAAAAAAGCIACACCATCACCATCACCATTAAGTCGA 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-600-783-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ADDRESSEE: KIM, Dae Young
STREET: Sosa Jukong Apt. 108-202, Sosa Bon-dong,
STREET: Sosa-ku
                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Identity: 55.556
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Align seg 1/1 to: US-09-101-135-8 from: 1 to: 272
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CITY: Kwangmyung-shi
STATE: Kyungki-do
COUNTRY: Republic of Korea
ZIP: 423-030
ADDRESSEE: CHANG, Seung Gu
STREET: Hyundal Apt. 71-203, Apkujong-dong,
STREET: Kangnam-ku
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Garden Heights Apt. 202-801, #100,
Hwangkeum-dong, Soosung-ku
                                                                                                                                                                                                                                                                                                                                                                                  Length:
REGISTRATION NUMBER: 36,663
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             alignment_block:
US-09-674-779-4/rev x US-09-101-135-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: 255
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: KIM, Chong Suhl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COUNTRY: Republic of Korea ZIP: 135-110
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY: Republic of Korea
                                                                                                                                                                              LENGIH: 272 amino acids
                                                                                                         TELEFAX: 703-816-4100 INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                          4.393
                                                                                                                                                                                                                                                                                                                                                                                  61.50
                                                                                                                                                                                                                            single
                                                                                                                                                                                                                                                                    MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bucheon-shi
Kyungki-do
                                                                                                                                                                                                    TYPE: amino acid
STRANDEDNESS: sir
                                                                                                                                                                                                                                                 linear
                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                               alignment_scores:
Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Seoul
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         422-230
                                                                                                                                                                                                                                                                                                                                                                                                             Ratio:
```

```
GENERAL INFORMATION:
APPLICANT: SONG, HO Yeong
APPLICANT: ROTHE, Mike
TITLE OF INVENTION: TRAFZ-Associated Kinase
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, 34th floor
                            COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk, 3.5 inch, 1.44MB storage COMPUTER: IBM PC/AT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-08-677-862-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps: 0
Percent Identity: 72.727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Align seg 1/1 to: US-08-600-783-13 from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              51 CCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 ProSerAspLysProHisHisHisHisHis 11
                                                                                                OPERATING SYSTEM: MS-DOS
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/600,783
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: US/08/677,862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-09-674-779-4/rev x US-08-600-783-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1 0
                                                                                                                                                                                                                            PRIOR APPLICATION DATA:
APPLICATION NUMBER: RR 95-2751
FILLING DATE: 15-FEB-1995
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequence 2, Application US/08677862
Patent No. 5874230
                                                                                                                                                                                                                                                                                                             NAME: Shahan Islam
REGISTRATION NUMBER: 32,507
REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                         TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-1000
TELEFAX: (212) 953-7249
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E: Floppy disk
IBM PC compatible
  Republic of Korea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LENGTH: 17 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Brezner, David J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5.800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ; MOLECULE TYPE: protein US-08-600-783-13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                    FILING DATE:
CLASSIFICATION: 514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CLASSIFICATION: 435
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                USA
                   706-040
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY:
COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
```

```
||||||| |||:::
590 HisValAlaProGlnGlnAsnAlaLeuHisHisHisHisGlyAsnSerSe 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ......AATAAGCT 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, 34th floor
CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     seq_name: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:US-09-252-571-2
                                                                                                                                                                                                                                                                                                                                                                                              Percent Identity: 43.478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFFWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/252,571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Align seg 1/1 to: US-08-677-862-2 from: 1 to: 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_documentation_block:
Sequence 2, Application US/09252571
Patent No. 5981250
GENERAL INFORMATION:
APPLICANT: SONG, Ho Yeong
APPLICANT: ROTHE, Mike
TITLE OF INVENTION: TRAF2-Associated Kinase
NUMBER OF SEQUENCES: 2
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEFAX: 415 38-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 763 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      T96-005/A63613
                                                                                                                                                                                                                                                                                                                                                                                                                                   alignment_block:
US-09-674-779-4/rev x US-08-677-862-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/677,862
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 CATGTCTTACCAAAA......
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               REFERENCE/DOCKET NUMBER: 79
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
                                                                                                                                                                       TYPE: amino acid
STRANDEDNESS: not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX: 415 398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            606 rHisHisHisHisHis 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 37 ACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1: 763 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                   56.50
4.708
52.174
                                                                                                                                                                                                                                         ; MOLECULE TYPE: peptide US-08-677-862-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
                                                                                                                                                                                                                  TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                   Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CA
                                                                                                                                                                                                                                                                                                                             alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE:
```

```
||||||| |||:::
590 HisValAlaProGlnGlnAsnAlaLeuHisHisHisHisGlyAsnSerSe 606
                                                                                                                                                                                                                                                                                                                             60 CATGTCTTACCAAAA.....AATAAGCT 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ADDRESSE: Flehr, Hohbach, Test, Albritton & Herbert STREET: 4 Embarcadero Center, 34th floor CITY: San Francisco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq_name: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:US-09-434-065-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23
                                                                                                                                          56.50 Length: 23
4.708 Gaps: 1
52.174 Percent Identity: 43.478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                        Align seg 1/1 to: US-09-252-571-2 from: 1 to: 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: SONG, HO Yeong
APPLICANT: ROTHE, Mike
TITLE OF INVENTION: TRAF2-Associated Kinase
CORRESPONDENCES: 2
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Brezner, David J
REGISTRATION NUMBER: 24,774
REFERENCE/DOCKET NUMBER: T96-005/A63613
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415 781-1989
TELEFAX: 415 398-3249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/09/434,065
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRIOR APPLICATION DATA: APPLICATION NUMBER: 08/677,862
                                                                                                                                                                                                                             alignment_block:
US-09-674-779-4/rev x US-09-252-571-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_documentation_block:
; Sequence 2. Application US/09434065
; Patent No. 6107074
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 not relevant
not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                       606 rHisHisHisHisHis 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            763 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                               37 ACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              56.50
                                    , MOLECULE TYPE: peptide US-09-252-571-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                     linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Quality:
Ratio:
                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRANDEDNESS:
STRANDEDNESS:
                                                                                                                                               Quality:
                                                                                                                                                                  Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 94111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            alignment_scores:
                                                                                                                           alignment_scores:
                   TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STATE:
```

```
GENERAL INFORMATION:
APPLICANT: Palleja, Estivill Xavier
APPLICANT: Pitchard, Melanie
APPLICANT: Pitchard, Melanie
APPLICANT: Vilato, Jordi Guimera
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
Patent No. 6251664
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Threonial Protein Kinase (MNB), Expressed in the
TITLE OF INVENTION: Neuronal Regions Affected in Down Syndrome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: Palleja, Estivill Xavier
APPLICANT: Palleja, Estivill Xavier
APPLICANT: Pitchard, Melanie
APPLICANT: Vilato, Jordi Guimera
TITLE OF INVENTION: Human Gene Sequence of the Down Syndrome Critical
Patent No. 6221664
TITLE OF INVENTION: Region of Human Chromosome 21, Coding for A Serine-
TITLE OF INVENTION: Threonine Protein Kinase (MNB), Expressed in the
TITLE OF INVENTION: Neuronal Regions Affected in Down Syndrome
TITLE OF INVENTION: NUMBER: US/08/789,275A
CURRENT APPLICATION NUMBER: US/08/789,275A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |||||| || |||:::
590 HisValAlaProGlnGlnAsnAlaLeuHisHisHisHisGlyAsnSerSe 506
                                                                                                                                                                                                                       990 HisValAlaProGlnGlnAsnAlaLeuHisHisHisHisGlyAsnSerSe 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 CATGTCTTACCAAAA.....AATAAGCT 38
                                                                                                                                                                                           60 CATGICITACCAAAA.....AATAAGCT 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  seq_name: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-789-275-5
                                                                                                                                                                                                                                                                                                                                                                                                         seq_name: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-789-275-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Identity: 43.478
Percent Identity: 43.478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   to: US-08-789-275-4 from: 1 to: 763
                                                                                                                                     to: 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps:
                                                                                                                                  Align seg 1/1 to: US-09-434-065-2 from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-674-779-4/rev x US-08-789-275-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   seq_documentation_block:
    Sequence 4, Application US/08789275A
    Patent No. 6251664
                                                        alignment_block:
US-09-674-779-4/rev x US-09-434-065-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              606 rHisHisHisHisHis 612
                                                                                                                                                                                                                                                                                                                                   606 rHisHisHisHisHis 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  37 ACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                         37 ACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56.50
4.708
52.174
Percent Similarity: 52.174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ratio:
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            aliqnment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH: 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-789-275-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Align seg 1/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQ ID NO 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURE:
```

```
590 HisvalalaProGlnGlnAsnAlaLeuHisHisHisGlyAsnSerSe 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .....AATAAGCT 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT: Conaway, Ronald C.
APPLICANT: Conaway, Joan W.
APPLICANT: Bradsher, John N.
TITLE OF INVENTION: RNA POlymerase Transcription Factor
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-524-757-30
                                                                                                                                                                                                                                                                                                                                                           Percent Identity: 43.478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     COMPUTER REALABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/524,757
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Align seg 1/1 to: US-08-789-275-5 from: 1 to: 763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE:

CLASSIFICATION 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION NUMBER: PCT/US94/13621

FILING DATE: 29-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/160087

FILING DATE: 30-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: HATE, John A.

REGISTRATION NUMBER: B35006CIPCIP

REFERENCE/DOCKET NUMBER: B35006CIPCIP

TELEPHONER: (214) 939-4500

TELEPHONER: (214) 939-4500

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 10 amino acids:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS STREET: 1201 Elm Street, Suite 4500
                        CURRENT APPLICATION NUMBER: US/08/789,275A CURRENT FILING DATE: 1997-01-28 NUMBER OF SEQ ID NOS: 6
                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-674-779-4/rev x US-08-789-275-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Sequence 30, Application US/08524757
; Patent No. 5792634
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          606 rHisHisHisHisHis 612
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  37 ACACCATCACCATCACCAT 19
FILE REFERENCE: U 011114-4
                                                                                                                                                                                                                                                                                                                                   Ratio: 4.708
Percent Similarity: 52.174
                                                                                                                                        TYPE: PRT
ORGANISM: Rat norvegicus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      60 CATGTCTTACCAAAA...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq_documentation_block;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COUNTRY: US
ZIP: 75270-2197
                                                                                                                                                                                                                                                                                 alignment_scores:
Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Dallas
                                                                                                             LENGTH: 763
                                                                                                                                                                                                                                                                                                                                                                                                      alignment_block:
                                                                                                                                                                                                           US-08-789-275-5
                                                                                              SEQ ID NO 5
```

```
Sequence 31, Application US/08524757
Patent No. 5792634
GENERAL INFORMATION:
GENERAL INFORMATION:
GENERAL CANT:
CONMANY, Ronald C.
APPLICANT:
CONMANY, Joan W.
APPLICANT:
Bradeher, Joan W.
TITLE OF INVENTION: RNA POLYMerase Transcription Factor
CORRESPONDENCE ADDRESS:
ADDRESSEE: RICHARDS, MEDLOCK & ANDREWS
STREET: 1201 Elm Street, Suite 4500
                                                                                                                                                                                                                                                                                                                                                                                                                   seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-524-757-31
                                                                                                                                                                                                 Percent Identity: 80.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CURRENT APPLICATION BATE:

APPLICATION NUMBER: US/08/524,757
                                                                                                                                                                                                                                                                                             Align seg 1/1 to: US-08-524-757-30 from: 1 to: 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REGISTRATION NUMBER: 37,345
REFERENCE/DOCKET NUMBER: B35006CIPCIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (214) 939-4500
TELEFAX: (214) 939-4600
INFORMATION FOR SEQ ID NO: 31:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/13621
FILING DATE: 29-NOV-1994
PRIOR APPLICATION NUMBER: US 08/160087
FILING DATE: 30-NOV-1993
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                     39 CTACACCATCACCATCACCATTAAGTCGAC 10
                                                                                                                                                                                                                                                                                                                                                            alignment_block:
US-09-674-779-4/rev x US-08-524-757-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: US
ZIP: 75270-2197
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                            6.222
90.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS:
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-524-757-30
                                                                                                                                                           56.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: Harre, John A. REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE: peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                           seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: amino acid
STRANDEDNESS: si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                linear
                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                           Quality:
                                                                                                                                                                            Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Dallas
                                                                                                                                    alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-08-524-757-31
```

Quality:

alignment\_scores:

Ratio: 6.222 Gaps: 0 Percent Similarity: 90.000 Percent Identity: 80.000

alignment\_block: US-09-674-779-4/rev x US-08-524-757-31 Align seg 1/1 to: US-08-524-757-31 from: 1 to: 10

```
out_format : pfs
OM of: US-09-674-779-4 to: PIR_71:*
```

Date: Jul 30, 2002 4:25 PM

About: Results were produced by the GenCore software, version 4.5, Copyright (c) 1993-2000 Compugen Ltd

Command line parameters:

Database: PIR\_71:\*
Database sequences: 283138
Database length: 96089334
Search time (sec): 28.830000 Search information block: Query: US-09-674-779-4 Query length: 60

Documentation.

1 hypothetical protein F28A21.23

1 hypothetical protein YOR134w - 1 selemprotein P precursor [val]

2 methyl CpG binding protein 2 - 1 methyl-CpG-binding protein 2 - 1 methyl-CpG-binding protein 2 - 15 hypothetical protein M50E8A.g transcription activator MafB - 1 hypothetical protein DKF2A34D1 hypothetical protein F1P2.220 otxl protein otxl - numan otxl protein - mouse homeodomain protein otxl - rat transcription intiation factor transcription factor oct-6 - hu transcription factor oct-6 - mouse the protein model. Sulvail phospholipase C (EC 3.1.4.3), phospholipase C (EC 3.1.4.3), phospholipase C (EC 3.1.4.3), calcium channel Br.1 - rabbit hypothetical protein F2413.210 neuroblast proliferation inhibition of the protein rat transforming protein (cbl) - hu (MBS1 protein - smut fungus (Us in Myothetical protein H06001.2 protein R1504.13 [imported] - calcium channel protein alpha-in hypothetical protein f04010.4 proline-rich protein precursor transcription factor-kr - mouse homeotic protein otxl - human Down-syndrome-critical region p hypothetical protein 72884.4 -polycomb (Pc) protein fruit f PBF-G2 (HFK-2) protein - fruit f transcription factor HFK1 - human transcription factor Brn-1 - mc hypothetical protein DKFZp434E1 SynGAP-bl protein - rat shs protein - fruit fly (Drosd Documentation 10057 4406 4406 4406 4406 4609 167.77 163.84 163.84 115.08 1143.79 1151.08 1151.08 1154.07 1140.07 1140.07 1140.08 1136.68 1136.68 1136.68 1148.28 1146.88 1146.88 1146.88 1146.88 1146.88 1146.88 1146.88 1146.88 1146.88 1146.88 1146.88 1146.88 62.50 62.50 63.60 64.00 65.60 65 Strd pirl:A56018 pirl:S30205 pirl:A40168 pir2:S27473 pir2:T23056 pir2:149529 pir2:S39406 score\_list: pir2:T04874 pir2:S60988 pir1:OMRTSP pir2:T31611 pir2:JC4898 pir1:A38565 :137451 pir2:A43817 pir2:T24446 pir2:S23737 : A49077 pir2:JH0672 pir2:T13804 pir2:146477 pir2:A40721 pir2

```
| Ras-GTPase activating prote
| ras GTPase-activating prote
                      | ras GTPase-activating prote | Nature 1 class homeodomain | calcium channel protein alp | hypothetical protein YKL111c
                                                                                                                                                                                 hypothetical professions that are cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C;Accession: T04874
R;Bevan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; May submitted to the Protein Sequence Database, February 1999
A;Accession: T04874
                                                                                                                                                                                                                                                                                                                                                                                                  A;Cross-references: EMBL:AL035526
A;Experimental source: cultivar Columbia; BAC clone F28A21
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Identity: 72.727
  1.79
1.79
2.14
2.17
2.57
134.06
133.79
145.20
128.11
150.15
54.00
54.00
53.50
53.50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                A;Map position: 4
A;Introns: 680/3; 754/1; 881/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            alignment_block:
US-09-674-779-4/rev x T04874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ratio: 5.818
Percent Similarity: 100.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               64.00
                                                                                                                                                                    seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                              A; Residues: 1-1057 <BEV>
                                                                                                                            seq_name: pir2:T04874
                                                                                                                                                                                                                                                                                                                                                          A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Note: F28A21.230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           alignment_scores:
                                     pir2:T06387
pir2:A41098
pir2:T14270
                   pir2:T14259
                                                                               pir2:S37939
```

928 ProArgAsnAsnValHisHisHisHisHis 938 Align seg 1/1 to: T04874 from: 1 to: 1057 51 CCAAAAATAAGCTACACCATCACCATCACCAT 19

seq\_name: pir2:S60988

hypothetical protein YOR134w - yeast (Saccharomyces cerevisiae)
NAlternate names: hypothetical protein 03320; hypothetical protein YOR3320w
C;Species: Saccharomyces cerevisiae
C;Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 07-May-1999
C;Accession: \$60988; \$51690; \$67019; \$63865
R;Wiemann, S:; Rechmann, S.; Benes, V.; Voss, H.; Schwager, C.; Vlcek, C.; Stegemann, submitted to the EMBL Data Library, August 1995
A;Pescription: Sequencing of 51 kilobases on the right arm of chromosome XV from S. C.
A;Accession: \$60988 seq\_documentation\_block:

A. Wolecule type: DNA
A. Residues: 1-409 <WIE>
A. Crossreferences: EMBL: X90518; NID:g1050808; PID:g1050814
A. Crossreferences: EMBL: Rechmen, S.; Teodoru, C.; Banrevi, A.; Sander, C.; Valen
Submitted to the EMBL Data Library, December 1995
A. Perscription: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromos
A. Reference number: S61643
A. Reference number: S61643
A. Reference number: S61643

A; Molecule type: DNA A; Residues: 1-409 CBEN> A; Cross-references: BMBL:X94335; NID:g1262139; PID:e217752; PID:g1164978 R; Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansorge, submitted to the Protein Sequence Database, July 1996 A; Reference number: S66965

3

A; Accession: S67019

A; Molecule type: DNA

shs

A; Residues: 1-409 <VOS> A; Cross references: EMBL: 275042; NID: 91420343; PID: e252026; PID: 91420344; MIPS: YOR134 A; Experimental source: strain S288C

```
Align seg 1/1 to: S57963 from: 1 to: 476
                                                                                         Align seg 1/1 to: OMRTSP from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4.917
80.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.917
80.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                alignment_block:
US-09-674-779-4/rev x S57963
      alignment_block:
US-09-674-779-4/rev x OMRTSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          59.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           59.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Reference number: S57963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     seq_documentation_block:
                                                                                                                                                                                                                                                                                                                             seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Molecule type: mRNA
A; Residues: 1-476 <DES>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ratio:
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A, Accession: S57963
A, Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq_name: pir2:A41907
                                                                                                                                                                                                                                                                   seq_name: pir2:S57963
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Noclecule type: protein
A; Residues: 20-41;267-287;316-327 <HI2>
B; Residues: 20-41;267-287;316-327 <HI2>
B; Richittum, H.S.; Himeno, S.; HILL, K.E.; Burk, R.F.
Arch. Biochem. Blophys: 325, 124-128, 1996
A; Title: Multiple forms of selenoprotein P in rat plasma.
A; Reference number: $68322; MUID:96140605
A; Reference number: $68322; MUID:96140605
A; Residues: 20-27 <CHI>
C; Superfamily: selenoprotein P
C; Superfamily: selenoprotein P
C; Reywords: extracellular protein; predicted <SIG>
F; 1-19/Domain: signal sequence #status predicted <SIG>
F; 20-385/Product: selenoprotein P #status experimental <AMT>
F; 59,264,335,377,371,373,380,382/Modified site: selenocysteine #status predicted
F; 83,174,188,370,375/Binding site: carbohydrate (Asn) (covalent) #status predicted
F; 282,323/Modified site: selenocysteine #status experimental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Selenoprotein P precursor [validated] - rat
C; Species: Rattus norvegicus (Norway rat)
C; Species: Rattus norvegicus (Norway rat)
C; Accession: A0380; B40380; B40380; B40380; B60322
C; Accession: A0380; B40380; B60320;
Biol. Chem. 266, 10050-10053, 1991
A; Title: The cDNA for rat selenoprotein P contains 10 TGA codons in the open reading fra A; Accession: A40380; MUID:91244760
A; Reference number: A40380; MUID:91244760
A; Accession: A40380
A; Molecule type: mRNA
A; Residues: 1-385 < HIL>
A; Cross-references: GB:M63574; NID:9206893; PIDN:AAA42129.1; PID:9206894
A; Accession: B40380
R:Wiemann, S.; Rechmann, S.; Benes, V.; Voss, H.; Schwager, C.; Vlcek, C.; Stegemann, J. Yeast 12, 281-288, 1996
A;Title: Sequencing and analysis of 51 kb on the right arm of chromosome XV from Sacchar A; Reference number: $63860; MUID:97060020
A; Reference number: $63865
A; Status: nucleic acid sequence not shown; translation not shown
                                                                                                                                                                                                                                          A;Cross-references: EMBL:X90518; NID:91050808; PID:91050814
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 302 HisMetLeuProArgSerArgAlaLeuSerAspSerAsnAsnPheThrIl 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 318 eHisHisHisHisHisHisHisAlaLeuPheProSerPro 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Identity: 75.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps: 1
Percent Identity: 35.484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               36 ......CACCATCACCATCACCATTAAGTCGACCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                   A)Gene: SGD:BAG7
A)Cross-references: SGD:S0005660; MIPS:YOR134w
A)Map position: 15R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Align seg 1/1 to: S60988 from: 1 to: 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   60 CATGTCTTACCAAAAAATAAGCTA..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61.00
6.100
83.333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3.906
51.613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        alignment_block:
US-09-674-779-4/rev x S60988
                                                                                                                                                                                                                    A; Residues: 1-409 <WIW>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          seq_name: pir1:OMRTSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Ratio:
                                                                                                                                                                                               A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          alignment_scores
                                                                                                                                                                                                                                                                                                                 C; Genetics
```

```
methyl-tog-binding protein 2 - rat
Methyl-tog-binding protein 2 - rat
M.Alternate names: chromosomal protein MeCP2
C;Species: Rattus norvegicus (Norway rat)
C;Decies: Rattus norvegicus (Norway rat)
C;Decession: A41907; S41461
R;Lewis, J.D.; Meehan, R.R.; Henzel, W.J.; Maurer-Fogy, I.; Jeppesen, P.; Klein, F.;
Cell 69, 905-914, 1992
A;Title: Purification, aquence, and cellular localization of a novel chromosomal pro
A;Title: Purification, A41907; MUD:92298389
A;Status: preliminary
A;Molecule type: MRNA
A;Residues: 1-492 <LEW>A;Molecule type: MRNA
A;Residues: MR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Crossreferences: GB:M94064; NID:g205361; PIDN:AAA41584.1; PID:g205362
R;Nan, X.; Meehan, R.R.; Bird, A.
Nucleic Acids Res. 21, 4886-4892, 1993
A;Title: Dissection of the methyl-CpG binding domain from the chromosomal protein MeC
A;Retence number: S41461; MID:9432813
A;Contents: annotation; methyl CpG-binding domain
C;Keywords: chromosomal protein; DNA binding
F;78-162/Domain: methyl-CpG-binding #status experimental <MCG>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C;Species: Homo sapiens (man)
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 05-Nov-1999
C;Accession: 557963
R;d'Esposito, M.; Quaderi, N.A.; Ciccodicola, A.; Bruni, P.; Esposito, T.; D'Urso, M. submitted to the EMBL Data Library, July 1995
A;Description: Physical mapping and expression analysis of an X-linked gene encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Cross-references: EMBL: X89430; NID: 9899295; PIDN: CAA61599.1; PID: 9899296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    51 CCAAAAAATAAGCTACACCATCACCATCACCATTAAGTCGACCCT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Identity: 60.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length: 15
Gaps: 0
Percent Identity: 60.000
                                                               238 LeuProProSerGlyLeuHisHisHisHisHis 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
54 TTACCAAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                            methyl CpG binding protein 2 - human (fragment)
```

alignment\_block:

```
C. Accession: T45727 C. C. Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Submitted to the Protein Sequence Database, November 1999
A. Reference number: 223010
                                                                                                                                                                                                                       seq_documentation_block:
hypothetical protein DKFzp434D1319.1 - human
C;Species: Homo sapiens (man)
C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
C;Accession: T46375
R;Ottenwaelder, B;Obermaier, B;Mewes, H.W.; Gassenhuber, J.; Wiemann, S. submitted to the Protein sequence Database, January 2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 hypothetical protein FIP2.220 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Feb-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            51 CCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGACCCTCTGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Experimental source: adult testis; clone DKF2p434D1319 C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Experimental source: cultivar Columbia; BAC clone F1P2 C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Identity: 52.941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Identity: 56.250
                                                                                                               Length:
                                                                                    36 CACCATCACCATCACCATTAAGTCGACCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Align seg 1/1 to: T46375 from: 1 to: 439
                                          to: A56235 from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Cross-references: EMBL: AL137724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Cross-references: EMBL: AL132955
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alignment_block:
US-09-674-779-4/rev x T46375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       4.750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    alignment_block:
US-09-674-779-4/rev x T45727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           56.50
4.346
81.250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 57.00
                                                                                                                                                                                                                                                                                                                                                                                                A; Reference number: Z23031
A; Accession: T46375
                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-439 <AAA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-240 <CHO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Note: DKFZp434D1319.1
                                                                                                                                                                                  seq_name: pir2:T46375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   seq_name: pir2:T45727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ratio:
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A; Accession: T45727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Note: F1P2.220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Map position:
                                          Align seg 1/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       91 0 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 T 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                seq_documentation_block:

transcription activator MafB - chicken

C;Species allus gallus (chicken)

C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 20-Jun-2000

C;Accession: A56235

R;Kataoka, K.; Fujiwara, K.T.; Noda, M.; Nishizawa, M.

Mol. Cell. Biol. 14, 7581-7591, 1994

A;Title: MafB, a new Maf family transcription activator that can associate with Maf and A;Accession: A56235

A;Stetus: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A;Residues: 1-1585 <WIL>
A;Cross-references: EMBL:AL117200; NID:e1549770; PIDN:CAB55050.1; CESP:Y50E8A.g
A;Experimental source: clone Y50E8A
C;Genetics:
                                                                                                                                                                                                                                                          hypothetical protein Y50E8A, g - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-0ct-1999 #sequence_revision 29-0ct-1999 #text_change 29-0ct-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A;Cross-references: GB:D28600; NID:g516723; PIDN:BAA05938.1; PID:g516724
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Gene: CESP:Y50E8A.g
A;Introns: 25/3; 60/1; 133/2; 217/3; 270/3; 337/2; 400/1; 746/2
                                                                                                                         362 ProLysLysGluHisHisHisHisHisHisAlaGluSerPro 376
                                                                                              51 CCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGACCCT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Identity: 66.667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Identity: 66.667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A;Introns: #status absent
C;Superfamily: maf transforming protein; maf homology
C;Keywords: DNA binding; homodimer; leucine zipper
F;200-289/Domain: maf homology <MAF>
                                                                                                                                                                                                                                                                                                                                                                                                                                                    A;Status: preliminary; translated from GB/EMBL/DDBJA;Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                   Risteward, C. submitted to the EMBL Data Library, September 1999 A:Reference number: 221047
A:Reference number: 131611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          54 TTACCAAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Align seg 1/1 to: T31611 from: 1 to: 1585
                                               Align seg 1/1 to: A41907 from: 1 to: 492
US-09-674-779-4/rev x A41907
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-674-779-4/rev x T31611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                83.333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5.700
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     alignment_block:
US-09-674-779-4/rev x A56235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 58.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 57.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5.800
                                                                                                                                                                                                                                         seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type: DNA
A; Residues: 1-311 <KAT>
                                                                                                                                                                                            seq_name: pir2:T31611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          seq_name: pir2:A56235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                  C; Accession: T31611
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    C; Genetics:
```

```
seq_documentation_block:
HBF-G2 (HFK-2) protein - human
C; Species: Homo sapiens (man)
C; Species: Homo sapiens (man)
C; Species: Homo sapiens (man)
C; Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 24-Sep-1999
C; Accession: 137451
R; Wiese, S.; Murphy, D.B.; Schlung, A.; Burfeind, P.; Schmundt, D.; Schnulle, V.; Mat Biochim. Biophys. Acta 1262, 105-112, 1995
A; Title: The genes for human brain factor 1 and 2, members of the fork head gene fami A; Reference number: 137451; MUID:95322450
A; Accession: 137451
A; Molecule type: DNA

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       seq_documentation_block:

polycomb (Pc) protein - fruit fly (Drosophila melanogaster)

C.Species: Drosophila melanogaster

C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C.Accession: A3856

R.Paro, R.; Hogness, D.S.

R.Paro, R.; Hogness, D.S.

R.Paro, R.; Hogness, D.S.

A.Reference number: A38565; MUID:91095442

A.Reference number: A38565; MUID:91095442

A.Rocession: A38565

A.Status: preliminary

A.Molecule type: DNA

A.Rocesidues: 1-390 < PAR>
A.Cross-references: GB:X55702; NID:98321; PIDN:CAA39229.1; PID:9603986

C.Genetics: Planse: Pl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Gene: HBF-G2; HFK-2
C; Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F;162-253/Domain: fork head DNA-binding domain homology <FHD>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Identity: 52.941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Identity: 80.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A;Cross-references: FlyBase:FBgn0003042
C;Superfamily: polycomb protein; chromobox homology
F;26-63/Domain: chromobox homology <CBH>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                         to: T32443 from: 1 to: 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Align seg 1/1 to: A38565 from: 1 to: 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               130 LysGluLysLysHisHisHisHisHis 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   48 AAAAATAAGCTACACCATCACCAT 19
                                                                                                                                                                                                                                                                                  45 AATAAGCTACACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     56.00
5.091
64.706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56.00
6.222
90.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-674-779-4/rev x A38565
alignment_block:
US-09-674-779-4/rev x T32443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          seq_name: pir2:137451
                                                                                                                                                                                                                                                                                                                                                                                                                                     seq_name: pir1:A38565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               alignment_block:
                                                                                                                                              Align seg 1/1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Species: Homo saplens (man)
C; Species: Homo saplens (man)
C; Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jul-2000
C; Accession: JC4898
R; Shindoh, U: Kudoh, J: Maeda, H: Yamaki, A.; Minoshima, S:; Shimizu, Y.; Shimizu, N. Biochem. Biophys. Res. Commun. 225, 92-99, 1996
A; Title: Cloning of a human homolog of the Drosophila minibrain/rat dyrk gene from "the A; Reference number: JC4898; MUID:96332410
A; Accession: JC4898
A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Cross-references: DDBJ:D85759; NID:91526445; PIDN:BAA12866.1; PID:91526446
C; Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homold
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              C; Accession: T32443
R; Wilson, R.; Graco, T.; Sansone, J.
submitted to the EMBL Data Library, September 1997
A; Description: The sequence of C. elegans cosmid T28B4.
A; Reference number: 221168
A; Accession: T32443
A; Accession: T32443
A; Accession: T32443
A; Accession: T32443
A; Caros: reference bush
A; Residues: 1-219 <WILL>
A; Residues: 1-219 <WILL>
A; Cross: references: EMBL: AF026206; PIDN: AAB71260.1; GSPDB: GN00028; CESP: T28B4.4
A; Experimental source: strain Bristol N2; clone T28B4
C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            seq_documentation_block:
hyporthetical protein 72884.4 - Caenorhabditis elegans
hyporthetical protein 72886.6 - C. Species: Caenorhabditis elegans
C; Species: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 18-Feb-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               581 HisValalaProGlnGlnAsnAlaLeuHisHisHisHisGlyAsnSerSe 597
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 CATGTCTTACCAAAA.....AATAAGCT 38
                                                                                                                                                                                                                                         .....SerProLeuPro 29
                                                                                                                                                        48 AAAAATAAGCTACACCATCACCATCACCATTAAGTCGACCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Identity: 77.778
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Down-syndrome-critical-region protein - human N; Alternate names: Drosophila minibrain protein homolog
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Percent Identity: 43.478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Align seg 1/1 to: JC4898 from: 1 to: 754
                                                                                                                                                                                                           to: T45727 from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        597 rHisHisHisHisHis 603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          37 ACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Map position: X
A; Introns: 34/2; 138/2; 184/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 4.708
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-09-674-779-4/rev x JC4898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ratio: 6.222
Percent Similarity: 100.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              56.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                          documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Molecule type: DNA A; Residues: 1-754 <SHI>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            seq_name: pir2:T32443
                                                                                                                                                                                                                                                                                                                                                    seq_name: pir2:JC4898
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Gene: CESP: T28B4.4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         alignment_block:
                                                                  Align seg 1/1
```

```
F;377-425/Domain: RING finger homology <RRN>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C. Species: Homo sapiens (man)
C. Species: Homo sapiens (man)
C. Species: Homo sapiens (man)
C. Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 02-Sep-2000
C. Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 02-Sep-2000
C. Accession: Asiant M.; Morse III, H.C.; Langdon, W.Y.
C. Splake, T.J.; Shapiro, M.; Morse III, H.C.; Langdon, W.Y.
C. Splake, T.J.; Shapiro, M.; Morse III, H.C.; Langdon, W.Y.
C. Title: The sequences of the human and mouse c-cbl proto-oncogenes show v-cbl was gener A. Title: The sequences of the human and mouse c-cbl proto-oncogenes show v-cbl was gener A. Accession: A43817
A. Status: preliminary
A. Status: preliminary
A. Status: preliminary
A. Status: BRBL: A57110; NID:929730; PIDN:CAA40393.1; PID:929731
C. Superfamily: RING finger homology
C. Superfamily: RING finger homology
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         C;Species: Homo saptens (man)
C;Species: Homo saptens (man)
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 12-Sep-1997
C;Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 12-Sep-1997
C;Accession: A34743
R;Murphy, D.B.; Wises, S.; Burfeind, P.; Schmundt, D.; Mattei, M.G.; Schulz-Schaeffer, Genomics 21, 551-557, 1994
A;Title: Human brain factor 1, a new member of the fork head gene family.
A;Reference number: A54743; MUID:95048332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A;Cross-references: GDB:433550
A;Map position: 14q12-14q12
C;Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology
F;169-260/Domain: fork head DNA-binding domain homology <FHD>
                                                                                                                                                                                              45 ProGlaHisHisHisHisHisHisHisHisHisProProProPr 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   51 CCAAAAATAAGCTACACCATCACCATTAACTTAAGTCGACCTTTGCC 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 CCAAAAAATAAGCTACACCATCACCATCACCATTAAGTCGACCCTCTGCC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps: 0
Percent Identity: 52.941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Status: preliminary; nucleic acid sequence not shown A;Molecule type: mRNA
A;Residues: 1-476 <MUR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
                                                                                                    to: I37451 from: 1 to: 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Align seg 1/1 to: A54743 from: 1 to: 476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               seq_documentation_block:
transforming protein (cbl) - human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       seq_documentation_block:
transcription factor HFK1 - human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  56.00
5.091
64.706
alignment_block:
US-09-674-779-4/rev x 137451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Gene: GDB:FKHL4; HBF-1; HFK1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: GB:X74142
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           alignment_block:
US-09-674-779-4/rev x A54743
                                                                                                                                                                                                                                                                                                                                                                                                                       seq_name: pir2:A54743
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Quality:
Ratio:
Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_name: pir2:A43817
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        alignment_scores:
                                                                                                    Align seg 1/1
                                                                                                                                                                                                                                                                                                                                                           61 0 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61 0 61
                                                                                                                                                                                                                                                                                            1 T 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1 T 1
```

```
Length: 12
Gaps: 0
Percent Identity: 66.667
                                                                                                                                                   36 CACCATCACCATCACCATTAAGTCGACCCTCTGCCT 1
                                                                                                          Align seg 1/1 to: A43817 from: 1
            56.00
5.600
83.333
                                                                  alignment_block:
US-09-674-779-4/rev x A43817
alignment_scores:
Quality:
Ratio:
Percent Similarity:
```

```
OM of: US-09-674-779-4 to: SwissProt_40:*
```

Date: Jul 30, 2002 4:31 PM

About: Results were produced by the GenCore software, version Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-WODEL-frame\_fright model - DEV=xlp
-0=/cgn2\_1/GVSPTO\_spool/USO9674779/runat\_30072002\_151754\_7376/app\_query.fasta\_1.116
-0=/cgn2\_1/GVSPTO\_spool/USO9674779/runat\_30072002\_151754\_7376/app\_query.fasta\_1.116
-0B=Swissprot\_40 - GWTF=fastan - SDFETX=rsp - GAPOP=12.000
-GAPOP=4.500 - GWTF=Fastan - SDFETX=0.000
-GAPOP=4.500 - GGAPEXT=0.050 - KGAPOP=10.000 - KGAPEXT=0.500
-FGAPOP=4.500 - GGAPEXT=7.000 - YGAPOP=10.000 - YGAPEXT=0.500
-FGAPOP=6.000 - FGAPEXT=7.000 - YGAPOP=110.000
-YGAPEXT=0.500
-TRAN=human40.cdi -LIST-45 - DOCALIGN=200 - THR\_SCORE=pct
-TRAN=human40.cdi -LIST-45 - DOCALIGN=200 - THR\_SCORE=pct
-TRAN=human40.cdi -LIST-45 - DOCALIGN=200 - THR\_SCORE=pct
-THR\_MAX=100 - THR\_NIN=0 - ALIGN=12 - MOXIGN=000000000
-USER-USO9674779\_@CGGN1\_1\_45 - NCPU=6 - LCPU=3 - LOMELOG
-DEV\_TIMEOUT=120 - WARN\_TIMEOUT=30 - NO\_XLPXY - WAIT - THREADS=1

Query length: 60
Database: SwissProt 40:\*
Database sequences: 105224
Database length: 38719550
Search time (sec): 16.440000 Search information block: Query: US-09-674-779-4

2164 | P97445 mus musculus (mouse) 2212 | P54282 rattus norvegicus (rat 111 | P36074 saccharomyces cerevis homo sapiens (human 54.00 137.87 54.00 137.80 54.00 137.61 54.00 137.56 54.00 136.69 54.00 136.69 54.00 130.80 54.00 130.80 54.00 130.80 54.00 130.80 54.00 130.80 SwissProt\_40:CCAA\_HUMAN SwissProt\_40:CCAA\_MOUSE SwissProt\_40:CCAA\_RAT -SwissProt\_40:YKL1\_YEAST score\_list:

CODON, UGA.

```
dictyostelium disc
mus musculus (mous
mus musculus (mous
drosophila melanog
                                                    saccharomyces cere
                                                                                                                                                                                                     Eukaryota, Metazoa, Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                      Hill K.E., Lloyd R.S., Yang J.-G., Read R., Burk R.F., "The CDNA for rat selenoprotein P contains 10 TGA codons in the open reading frame.";
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN-WISTAR; TISSUE-Kidney;
MEDLINE-9516461; PubMed-7580;
Saijoh K., Saiton N., Lee M.J., Fujii M., Kobayashi T., Sumino K.;
Molecular cloning of cDNA encoding a bovine selenoprotein P-like protein containing 12 selenocysteines and a (His-Pro) rich domain
                          Q64733 n
P31368 c
P50089 s
                                                                                                                                                                                                                                                   SEQUENCE FROM N.A., AND SEQUENCE OF 20-41; 267-287 AND 316-327
                                                                                                                              01-MAY-1992 (Rel. 22, Created)
01-MAY-1992 (Rel. 22, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                  insertion, and its regional expression.";
Brain Res. Mol. Brain Res. 30:301-311(1995).
  137.00
136.90
135.96
133.28
                                                                                                                                                                                                                                                                                                                      Biol. Chem. 266:10050-10053(1991).
                                                                                                                                                                                                                                                                         MEDLINE=91244760; PubMed=2037562;
                                                                                                                                                                   Selenoprotein P precursor (SeP). SEPP1 OR SELP.
  53.00
53.00
53.00
53.00
                                                                     seq_name: SwissProt_40:SELP_RAT
                                                                                                           STANDARD;
                                                                                                                                                                                         Rattus norvegicus (Rat).
SwissProt_40:CAR2_DICDI -
SwissProt_40:SELP_MOUSE -
SwissProt_40:FALP_MOUSE -
SwissProt_40:PDMI_DROME -
SwissProt_40:PDMI_DROME -
SwissProt_40:YG51_YEAST -
                                                                                            seq_documentation_block:
                                                                                                                                                                                                                                                                 TISSUE=Liver;
                                                                                                          SELP_RAT P25236;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        REVIEW.
```

MEDILIRE-95017128; PubMed=7931697;

Burk R.F., Hill K.E.;

Burk R.F., Hill K.E.;

"Selenoprotein P. A selenoprotein P. A selenoprotein.";

J. Nutr. 124:1891-1897(1994).

-!- FUNCTION: MIGHT BE RESPONSTBLE FOR SOME OF THE EXTRACELLULAR ANTIOXIDANT DEFENSE PROPERTIES OF SELENIUM OR MIGHT BE INVOLVED IN THE TRANSPORT OF SELENIUM.
-!- SUBCELLULAR LOCATION: Extracellular. -1- MISCELLANEOUS: THE SELENOCYSTEINES ARE ALL ENCODED BY THE OPAL TISSUE SPECIFICITY: MADE IN THE LIVER AND SECRETED INTO THE PLASMA

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).

Selenium; Selenocysteine; Plasma. EMBL; M63574; AAA42129.1; -. EMBL; D25221; BAA04950.2; -. PIR; A40380; OMRTSP. Glycoprotein; Signal; St GIAIL 20 385 DOMAIN 244 252 SE\_CYS 264 264 SE\_CYS 282 282 SE\_CYS 282 282 SE\_CYS 333 333 SE\_CYS 355 SE\_CYS 355

146.61

53.00

```
Coy J.F., Sedlacek Z., Baechner D., Delius H., Poustka A.;
"A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3'-untranslated region of the methyl-cpc-binding protein 2 gene (MeCP2) suggests a regulatory role in dene expression ".
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE-97130625; PubMed-8976388; Vilain A., Apiou F., Vogt N., Dutrillaux B., Malfoy B.; Vilain A., Apiou F., Vogt N., Dutrillaux B., Malfoy B.; Assignment of the gene for methyl-CpG-binding protein 2 (MECP2) to human chromosome band Xq28 by in situ hybridization."; Cytogenet. Cell Genet. 74:293-294(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                  (POTENTIAL).
                                                                                                (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Reichwald K., Rosenthal A., Kloschis P., Platzer M.; "Mapping and sequence analysis of the human McCP2 locus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Thiesen J., Straetling W.H.; Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                            PSIGOR; O1523;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last saquence update)
Methyl-CpG-binding protein 2 (MeCP-2 protein) (MeCP2).
                                                                                                                             013BDF4FB741E3E8 CRC64;
                                                      (GLCNAC. . .) (GLCNAC. . .)
                                                                                                                                                                                     Length: 12
Gaps: 0
Percent Identity: 75.000
                                                                                                                                                                                                                                                                                                                                  238 LeuproProSerGlyLeuHisHisHisHisHis 249
                                                                                                                                                                                                                                                                                                                                                                                                                     486 AA.
                                                                                                                                                                                                                                                                                                                  54 TTACCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                       Align seg 1/1 to: SELP_RAT from: 1 to: 385
                                                      N-LINKED
N-LINKED
N-LINKED
N-LINKED
                                                                                                                 N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Jum. Mol. Genet. 8:1253-1262(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE=99299240; PubMed=10369871;
                                                                                                                             42614 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 10-486 FROM N.A. TISSUE-Skeletal muscle;
                                                                                                                                                                                                                                                                                                                                                                           seq_name: SwissProt_40:MEC2_HUMAN
                                                                                                                                                                                                                                                             US-09-674-779-4/rev x SELP_RAT
                                                                                                                                                                                      61.00
6.100
83.333
                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
             373
380
382
83
174
1188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human)
371
373
380
382
83
174
188
370
375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      in qene expression
                                                                                                                                                                                                                                                                                                                                                                                                        seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Fukuda M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TISSUE-Placenta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9606;
                                                                                                                                                                                                                   Percent Similarity:
                                                                                                                                                                                         Quality:
                                                                                                                                                                                                       Ratio:
                                                                                                                                                                            alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                       MEC2_HUMAN
                                                                                                                                                                                                                                               aliqnment_block:
SE_CYS
SE_CYS
SE_CYS
SE_CYS
CARBOHYD
                                                                      CARBOHYD
CARBOHYD
                                                                                                     CARBOHYD
                                                                                                                 CARBOHYD
                                                                                                                                SEOUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kudo
                                                                                                                                                                                                                                                                                                                                                                                                                          SHEFFFFFF
```

```
GENOME.

-I- TISSUE SPECIFICITY: PRESENT IN ALL ADULT SONATIC TISSUES TESTED.

-I- DISEASE: DEFECTS IN MECP2 ARE THE CAUSE OF RETT SYNDROME (RTT), AN X-LINKED DOMINANT DISEASE. RTT IS A PROGRESSIVE NEUROLOGIC DEVELOPMENTAL DISCABES. RTT IS A PROGRESSIVE NEUROLOGIC RETARDATION IN FEMALES. PATIENTS APPEAR TO DEVELOP NORMALLY UNTIL RETARDATION IN FEMALES. PATIENTS APPEAR TO DEVELOP NORMALLY UNTIL HAND MOVEMENTS AND DEVELOP MICROCEPHALY, SEIZURES, AUTISM, ATAXIA, INTERMITTENT HYPERVENTILATION AND STRECOTYPIC HAND MOVEMENTS.

AFTER INITIAL REGRESSION, THE CONDITION STABILIZES AND PATIENTS USUALLY SURVIVE INTO ADULTHOOD.

--- SIMILARITY: CONTAINS (R)GRP(K) MOTIFS AND SPKK MOTIFS THAT HAVE BEEN FOUND TO BIND TO MINOR GROOVE OF AT-RICH DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NEUTOIOGY 58:226-230(2002).

-i- FUNCTION: CHROMOSOMAL PROTEIN THAT BINDS TO METHYLATED DNA. IT CAN BIND SPECIFICALLY TO A SINGLE METHYL-CPG PAIR. IT IS NOT SIND SPECIFICALLY TO A SINGLE METHYL-CPGS. MEDIATES INFLUENCED BY SEQUENCES FLANKING THE METHYL-CPGS. MEDIATES TRANSCRIPTIONAL REPRESSION THROUGH INTERACTION WITH HISTONE DEACETYLASE AND THE COREPRESSOR SIN3A.

-i- SUBCELLULAR LOCATION: NUCLEAR. COLOCALIZED WITH METHYL-CPG IN THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Mutation analysis of the methyl-CpG binding protein 2 gene (MECP2) in patients with Rett syndrome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANTS RTT W-106; F-124; C-13; C-134; R-152; M-158 AND C-306. MEDLINE-20439334; PubMed-10991688; Cobata K., Matsuishi T., Yamashita Y., Fukuda T., Kuwajima K., Obata K., Matsuishi T., Nagamitsu S., Iwanaga R., Kimura A., Omori I., Endo S., Mori K., Kondo I.;
                                      D'Urso M., Brown S.D.M.; "Isolation, physical mapping, and Northern analysis of the X-linked human gene encoding methyl CpG-binding protein, MECP2.";
           Esposito M., Quaderi N.A., Ciccodicola A., Bruni P., Esposito T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=20465115; PubMed=11007980; Orrico A., Lam C., Galli L., Dotti M.T., Hayek G., Tong S.F., Poon P.M., Zappella M., Federico A., Sorrentino V.; "MECP2 mutation in male patients with non-specific X-linked mental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Dotti M.T., Orrico A., De Stefano N., Battisti C., Sicurelli F., Severi S., Lam C.W., Galli L., Sorrentino V., Federico A.; "A Rett syndrome MECP2 mutation that causes mental retardation in
                                                                                                                                                                                                                                                                                                                                                               "Rett syndrome is caused by mutations in X-linked MECP2, encoding
                                                                                                                                                                                                                                                                                                  Amir R.E., Van den Veyver I.B., Wan M., Tran C.Q., Francke U., Zoghbi H.Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANTS RTT R-101; W-106; M-158 AND C-306, AND VARIANT K-397
MEDLINE-20439335; PubMed-10991689;
                                                                                                                                                                  Reichwald K., Bauer D., Brenner V., Drescher B., Coy J.,
Kioschis P., Korn B., Nyakatura G., Platzer M., Poustka A.,
Sandoval N., Rosenthal A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hampson K., Woods C.G., Latif F., Webb T.; "Mutations in the MECP2 gene in a cohort of girls with Rett
                                                                                                                                                                                                                                   Submitted (DEC-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                              VARIANTS RTT TRP-106; CYS-133; SER-155 AND MET-158.
MEDLINE=96327611; PubMed=8672133;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          tients with Rett syndrome.";
Med. Genet. 37:608-610(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Med. Genet. 37:610-612(2000).
                                                                                                                                                                                                                                                                                                                                                                                        methyl-CpG-binding protein 2.";
Nat. Genet. 23:185-188(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 retardation.";
FEBS Lett. 481:285-288(2000).
                                                                                                         Mamm. Genome 7:533-535(1996).
                                                                                                                                                    SEQUENCE OF 10-486 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         VARIANT RTT VAL-140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT RTT VAL-140.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PubMed=11805248;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GENOME.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               syndrome.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial or still require a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                               \label{eq:first}  \text{/FTId=VAR\_010278.} \\ \text{FPECTS WOMEN AND MEN).} \\ \text{/FTId=VAR\_010279.} 
                                                                                                                                                                                                                                                 Transcription regulation; Repressor; DNA-binding; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PAVP -> RLC (IN REF. 6).
E -> G (IN REF. 3).
EB6A33233AEDA566 CRC64;
                                                                                                                                                                                                                                                                                                                                                             L -> F (IN RTT).
/FTId=VAR_010277.
                                                                                                                                                                                                                                                                                                                                        R -> W (IN RTT). /FTId=VAR_010272.
                                                                                                                                                                                                                                                                                                                                                                                           /FTId=VAR_010273.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              010274.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FTId=VAR_010282.
                                                                                                                                                                                                                                                                                                                                                                                                                                                         /FTId=VAR_010280.
                                                                                                                                                                                                                                                                                                                  P -> R (IN RTT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        T -> M (IN RTT). /FTId=VAR_010275.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FTId=VAR_010281.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /FTId=VAR 010283
                                                                                                                                                                                                                                                                                                                                                                                  -> C (IN RTT).
                                                                                                                                                                                                                                                                                                                                                                                                       -> C (IN RTT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                -> R (IN RTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -> S (IN RTT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -> C (IN RTT
                                                                                                                                                                                                                                                                                   POLY-ALA.
POLY-HIS.
POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              /FTId=VAR
                                                                                   EMBL, L37298, AAC32737.1, -.
EMBL, Y12643, CAA73190.1, -.
EMBL, X99686, CAA68001.1, -.
EMBL, AF030876, AAC08757.1, -.
EMBL, AF031078, AAC08758.1, -.
EMBL, AJ32917; CAB46446.1, -.
                                                                                                                                                                                                                                                                Polymorphism
                                                                                                                                                                                            InterPro; IPR000637; AT_hook.
InterPro; IPR001739; MBD.
Pfam; PF01429; MBD; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               486 AA; 52440 MW;
                                                                                                                                                    EMBL; X89430; CAA61599.1; -. EMBL; X94628; CAA64331.1; -.
                                                                                                                                                                                                                             SMART; SM00384; AT_hook; 1. SMART; SM00391; MBD; 1.
                                                                                                                                                                                                                                                                                             372
393
101
                                                                                                                                                                                                                                                                                                                                      901
                                                                                                                                                                                                                                                                                                                                                            124
                                                                                                                                                                                                                                                                                                                                                                                 133
                                                                                                                                                                                                                                                                                                                                                                                                     134
                                                                                                                                                                                                                                                                                                                                                                                                                          140
                                                                                                                                                                                                                                                                                                                                                                                                                                               152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      397
                                                                                                                                                                                                                                                             Disease mutation;
                                                                                                                                                                                                                                                                                            366
384
101
                                                                                                                                                                                                                                                                                                                                      106
                                                                                                                                                                                                                                                                                                                                                            124
                                                                                                                                                                                                                                                                                                                                                                                133
                                                                                                                                                                                                                                                                                                                                                                                                     134
                                                                                                                                                                                                                                                                                                                                                                                                                         140
                                                                                                                                                                                                                                                                                                                                                                                                                                             152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   155
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                306
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      397
                                                                                                                                                                          MIM; 300005;
MIM; 312750;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                 VARIANT
                                                                                                                                                                                                                                                                                                                                      VARIANT
                                                                                                                                                                                                                                                                                                                                                           VARIANT
                                                                                                                                                                                                                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                    VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                        VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                             VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                       DOMAIN
```

```
Percent Identity: 60.000
                Length:
                                                              alignment_block:
US-09-674-779-4/rev x MEC2_HUMAN
             59.00
4.917
                                        80.000
               Quality:
                                    Percent Similarity:
                           Ratio:
alignment_scores:
```

Align seg 1/1 to: MEC2\_HUMAN from: 1 to: 486

seq\_name: SwissProt\_40:MEC2\_RAT

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Burpen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Call 6995-914(1992).

CHOROSCOMAL PROTEIN THAT BINDS TO METHYLATED DNA. ";

CALL 69105-914(1992).

CALL 69207.

CALL 69207
                                                                                                                                                                                                                                                                 Bukaryota; Metazoa; Chórdata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- TISSUE SPECIFICITY: PRESENT IN ALL ADULT SOMATIC TISSUES TESTED.
-!- SIMILARITY: CONTAINS (R)GRP(K) MOTIFS AND SPKK MOTIFS THAT HAVE BEEN FOUND TO BIND TO MINOR GROOVE OF AT-RICH DNA.
-!- SIMILARITY: CONTAINS 1 METHYL-BINDING DOMAIN (MBD).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Transcription regulation; Repressor; DNA-binding; Nuclear protein. DOMAIN 96 149 MBD.
                                                                                                                                                                                                                                                                                                                                                                                                                      TISSUE=Brain;
MEDIINE=92298389; PubMed=1606614;
Lewis J.D., Meehan R.R., Henzel W.J., Maurer-Fogy I., Jeppesen
                                                                               01-APR-1993 (Rel. 25, Created)
01-APR-1993 (Rel. 25, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Methyl-CpG-binding protein 2 (MeCP-2 protein) (MeCP2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                451
53047 МW; A67E705C68BA2D38 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                362 ProLysLysGluHisHisHisHisHisHisHisAlaGluSerPro 376
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       51 CCAAAAATAAGCTACACCATCACCATTAAGTCGACCT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Identity: 60.000
                               492 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  484 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Align seg 1/1 to: MEC2_RAT from: 1 to: 492
                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-PRO.
                               PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; M94064; AAA41584.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      seq_name: SwissProt_40:MEC2_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     alignment_block:
US-09-674-779-4/rev x MEC2_RAT
                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PIR; A41907, A41907.
InterPro; IPR001739; MBD.
Pfam; PF01429; MBD; 1.
SMART; SM00391; MBD; 1.
                                                                                                                                                                                                                                       Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               59.00
4.917
80.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          443
492 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               seq_documentation_block:
ID MEC2_MOUSE STAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Klein F., Bird A.;
                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   alignment_scores
                            MEC2_RAT
Q00566;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
```

58.00

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions along as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Comparative analysis of the methyl CpG binding protein 2 locus in man and mouse reveals new untranslated sequences.";
submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
-! FUNCTION: CHROMOSOMAL PROTEIN THAT BINDS TO METHYLATED DNA. IT CAN BIND SPECIFICALLY TO A SINGLE METHYL-CPG PAIR. IT IS NOT INFLUENCED BY SEQUENCES FLARKING THE METHYL-CPGS. MEDIATES TRANSCRIPTIONAL REPRESSION THROUGH INTERACTION WITH HISTONE DEACETYLASE AND THE CORBRESSOR SINAA (BY SIMILARITY).
-! SUBCELLULAR LOCATION: NUCLEAR. COLOCALIZED WITH METHYL-CPG IN THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Coy J.F., Sedlacek Z., Baechner D., Delius H., Poustka A.; "A complex pattern of evolutionary conservation and alternative polyadenylation within the long 3'-untranslated region of the methyl-CpG-binding protein 2 gene (MeCP2) suggests a regulatory role in gene
                                                                                                                                                                                                                                                                                                                                                                                                                                             Hendrich B., Bird A.; "Identrication of a family of mammalian methyl-"Identification and characterization of a family of mammalian methyl-gcb binding proteins."; Mol. Cell. Biol. 18:6538-6547(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -i- SIMILARITY: CONTAINS (R)GRP(K) MOTIFS AND SPKK MOTIFS THAT HAVE BEEN FOUND TO BIND TO MINOR GROOVE OF AT-RICH DNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        regulation; Repressor; DNA-binding; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Reichwald K., Thiessen J., Wiehe T., Kioschis P., Straetling W.H., Rosenthal A., Platzer M.;
                                                                                                                                                                                                                 Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-THR.
62FD228F0118A49F CRC64;
                       16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Methyl-CpG-binding protein 2 (MeCP-2 protein) (MeCP2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           POLY-ALA.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLY - PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
MEDLINE=99299240; Pubmed=10369871;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Hum. Mol. Genet. 8:1253-1262(1999)
                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=98449942; Pubmed=9774669;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ΜW.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Interpro; IPR000637; AT_hook.
Interpro; IPR001739; MBD.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      52307
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRINTS; PR00929; ATHOOK. SMART; SM00391; MBD; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam; PF01429; MBD; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         366 33
384 33
440 4
                                                                                                                                                                                          Mus musculus (Mouse).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                             NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Transcription
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    expression.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
SOLUTION TO THE TERM TO THE TE
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PHOSPHORYLATION (AUTO-) (BY SIMILARITY). 89E17812A4900CD0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Petter R., Kwon-Chung K.J.;
"Disruption of the SNRI gene abolishes trehalose utilization in the pathogenic yeast Candida qlabrata.";
Infect. Immun. 64:5269-5273(1996).
-!- EUNCTION: ESSENTIAL FOR RELEASE FROM GLUCOSE REPRESSION. IT INTERACTS AND HAS FUNCTIONAL RELATIONSHIP TO THE REGULATORY PROTEIN SNR4. COULD PHOSPHORYLATES CAT8 (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -i - SUBCELLULAR LOCATION: ASSOCIATED WITH THE NUCLEAR MEMBRANE (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SIMILARITY).
-!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Candida glabrata (Yeast) (Torulopsis glabrata).
Eukaryota, Fungi, Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-NOV-1997 (Rel. 35, Created)
1-NOV-1997 (Rel. 35, Last sequence update)
15-JUL-1999 (Rel. 38, Last amotation update)
Carbon catabolite derepressing protein kinase (EC 2.7.1.-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       INEXT: YOUZIO; ZELLY
INTERPO: IPRO0219; Euk_pkinase.
InterPro; IPR002290; Ser_thr_pkinase.
InterPro; IPR002290; Ser_thr_pkinase.
Promosors Strk: 1.
PROSITE; PS00108; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
PROSITE; PS00108; PROTEIN_KINASE_DOM: 1.
Length: 11
Gaps: 0
Percent Identity: 72.727
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROTEIN KINASE.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                611 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BY SIMILARITY
                                                                                                                                                                                                                                                   Align seg 1/1 to: MEC2_MOUSE from: 1 to: 484
                                                                                                                                                                                                                                                                                                                                                     51 CCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=NCCLS84;
MEDLINE=97101049; PubMed=8945576;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  70049 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; L78130; AAB48642.1; -. HSSP; P00518; 2PHK.
                                                                                                                                                                                     US-09-674-779-4/rev x MEC2_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                              seq_name: SwissProt_40:SNF1_CANGA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                           5.800
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        53
68
161
194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SNF1 SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=5478;
                                               Ratio:
                                                                               Percent Similarity:
                                                                                                                                                        alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SNF1_CANGA
Q00372;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SO THE FEFFER SO TO SO TO SO THE SO T
```

alignment\_scores:

```
Guimera J., Casas C., Estivill X., Pritchard M.;
"Human minibrain homologue (MNBH/DYRK1): characterization, alternative
splicing, differential tissue expression, and overexpression in Down
syndrome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Song W.J., Sternberg L.R., Kasten-Sportes C., van Keuren M.L.,
Chung S.H., Slack A.C., Miller D.E., Glover T.W., Chiang P.W., Lou L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Isolation of human and murine homologues of the Drosophila minibrain gene: human homologue maps to 21q22.2 in the Down syndrome 'critical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Cloning of a human homolog of the Drosophila minibrain/rat Dyrk gene from 'the Down syndrome critical region' of chromosome 21."; Biochem. Biophys. Res. Commun. 225:92-99(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Guimera J., Casas C., Pucharcos C., Solans A., Domenech A., Planas A.M., Ashley J., Lovett M., Estivill X., Pritchard M.A.; "A human homologue of Drosophila minibrain (MNB) is expressed in the neuronal regions affected in Down syndrome and maps to the critical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDLINE-96332410; PubMed-8769099;
Shindoh N., Kudoh J., Maeda H., Yamaki A., Minoshima S., Shimizu Y.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DYRA_HUMAN 27ANDARD; PRT; 763 AA. 613627; Q92810; Q92582; Q9UNM5; Created) 16-NOV-1997 (Rel. 35, Created) 16-OCT-2001 (Rel. 40, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) 16-OCT-201 (Rel. 40, Last annotation update) 16-OC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A., VARIANTS F-415; H-681, AND ALTERNATIVE SPLICING MEDLINE=99263503; PubMed=10329007;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A. (ISOFORM 1).
Miki O., Masahira H., Yasufumi M., Naohiko S., Takahiro N.,
Hitoshi I., Nobuo N., Toshihiko E., Yoshiyuki S., Misao O.,
Ohira M., Seki N., Nagase T., Suzuki E., Nomura N., Ohara O.,
Hattori M., Sakaki Y., Eki T., Murakami Y., Saito T., Ichikawa H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Genomics 57:407-418(1999).
-1- FUNCTION: MAY PLAY A ROLE IN A SIGNALING PATHWAY REGULATING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Submitted (JAN-1997) to the EMBL/GenBank/DDBJ databases
             Gaps: 0
Percent Identity: 80.000
                                                                                                                                                                                                   611
                                                                                                                                                                                                   :
د
                                                                                                                                                                                                                                                    48 AAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                         2 GluAsnLysGluHisHisHisHisHis 11
                                                                                                                                                                               Align seg 1/1 to: SNF1_CANGA from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. (LONG ISOFORM).
MEDLINE=97131512; Pubmed=8975710;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. (LONG ISOFORM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Hum. Mol. Genet. 5:1305-1310(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=97026291; PubMed=8872470;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A. (ISOFORM 1).
                                                                                                                   US-09-674-779-4/rev x SNF1_CANGA
                                                                                                                                                                                                                                                                                                                                                                        seq_name: SwissProt_40:DYRA_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Genomics 38:331-339(1996).
6.333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                       seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TISSUE=Fetal brain;
                          Percent Similarity:
   Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Kurnit D.W.;
                                                                                            alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Shimizu N.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ohki M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         region,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               region.
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                              SUBCELLULAR LOCATION: Nuclear.
ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS; A LONG FORM (SHOWN HER), 1, 2, 3 AND 4; ARE PRODUCED BY ALTERNATIVE SPLICING.
TISSUE SPECIFICITY: UBJUUTIOUS. HIGHEST LEVELS IN SKELETAL MUSCLE,
TESTIS, FETAL LUNG AND FETAL KIDNEY.
DEVELOPMENTAL STAGE: EXPRESSED IN THE DEVELOPING CENTRAL NERVOUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SMART; SM00220; S_TKC; , PROPERLY KINASE_ATP; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase; ATP-binding; Nuclear protein; Phosphorylation; Alternative splicing;
NUCLEAR FUNCTIONS OF CELL PROLIFERARATION, PHOSPHORYLATES SERINES, THREONINES AND TYROSINES RESIDUS IN ITS SEQUENCE AND IN EXOGENOUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -> VEQHWMPGAFRM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SSHVVHLLVSP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PHOSPHORILATION (AUTO-) (BY SIMILARITY).
PHOSPHORILATION (AUTO-) (BY SIMILARITY).
PHOSPHORILATION (AUTO-) (BY SIMILARITY).
MISSING (IN ISOFORM 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 BIPARTITE NUCLEAR LOCALIZATION SIGNAL
                                                                                                                                                                                                                           PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES.
DISEASE: OVEREXPRESSED 1.5-FOLD IN FETAL DOWN SYNDROME BRAIN.
SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGSSGTSNSGRARS -> GASAISCSSWLVRH (IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AILRWSSTGCQVPLE (IN ISOFORM 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ROOFPAPLGWSGTEAPTQVTVETHPV ->
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TVSFTLEVHDVPV (IN ISOFORM 3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GGSSGTSNSGRARSDPTHQHRHSGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MISSING (IN ISOFORM 4).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MISSING (IN ISOFORM 2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MISSING (IN ISOFORM 3)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REF. 1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q -> H.
/FTId=VAR_009396.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            /FTId=VAR_009395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL).
PROTEIN KINASE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                BY SIMILARITY,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SER/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -> A (IN | -> P (IN | 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-SER.
POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ISOFORM 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro; IPR000719; Euk_pkinase.
InterPro; IPR002290; Ser_thr_pkinase.
PFam: PF00069; pkinase; Z.hr_pkinase.
SMART; SM00220; S_TKC; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Y -> F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; AF108830; AAD31169.1; -. HSSP; P27703; 1ERK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EMBL; U58496; AAC50939.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EMBL; U52373; AAB18639.1; -
EMBL; D85759; BAA12866.1; -.
EMBL; D86550; BAA13110.1; -.
                                                                                                                                                                                                                                                                                                 MNB/DYRK SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      321
321
78
529
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         763
584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           32
47
57
123
397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  681
                                                  SUBSTRATES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            321
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   559
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                681
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              32
47
57
123
397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MIM; 600855;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NP_BIND
BINDING
ACT_SITE
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
MOD_RES
MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MOD RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       /ARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                VARIANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DOMAIN
```

REF. 1).

FT SO

```
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its way non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SUBCELLULAR LOCATION, AND SUBUNIT.
MEDLINE-97224401; PubMed-9070862;
Song W.J., Chung S.H., Kurnit D.M.;
Song W.J., Chung S.H., Kurnit D.M.;
"The murine Dyrk protein maps to chromosome 16, localizes to the nucleus, and can form multimers.";
Biochem. Biophys. Res. Commun. 231:640-644(1997).
- FUNCTION: MAY PLAT A ROLE IN A SIGNALING PATHWAY REGULATING
- NUCLEAR FUNCTIONS OF CELL PROLIFERARATION, PHOSPHORYLAATES SERINES,
THREONINES AND TYROSINES RESIDUES IN ITS SEQUENCE AND IN EXOGENOUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Song W.J., Sternberg L.R., Kasten-Sportes C., van Keuren M.L.,
Chung S.H., Slack A.C., Miller D.E., Glover T.W., Chiang P.W., Lou L.,
Kurnit D.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Isolation of human and murine homologues of the Drosophila minibrain gene: human homologue maps to 21q22.2 in the Down syndrome 'critical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
16-CCT-2000 (Rel. 40, Last annotation update)
16-CT-2001 (Rel. 40, Last annotation update)
16-CT-2001 (Rel. 40, Last annotation regulated kinase 1A
16-CT-27.1.-) (Protein kinase minibrain homolog) (MNBH) (MP86) (Dual specificity YAK1-related kinase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -i- PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus
                                                                                                                                                                                                                                                                                                                                      590 HisValAlaProGlnGlnAsnAlaLeuHisHisHisHisGlyAsnSerSe 606
                                                                                                                                                                                                                                                                                                 60 CATGTCTTACCAAAA.....AATAAGCT 38
A -> G (IN REF. 1).
A -> P (IN REF. 1).
7C3A52A3CBB04FB5 CRC64;
                                                                                                                                                            Percent Identity: 43.478
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       763 AA
                                                                                                                    Length:
                                                                                                                                             Gaps:
                                                                                                                                                                                                                                                           to: 763
                                                                                                                                                                                                                                                           Align seg 1/1 to: DYRA_HUMAN from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBCELLULAR LOCATION: Nuclear.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=SWISS WEBSTER / NIH;
MEDLINE=97131512; Pubmed=8975710;
  592 592 A
679 679 A
763 AA; 85584 MW;
                                                                                                                                                                                                                        US-09-674-779-4/rev x DYRA_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq_name: SwissProt_40:DYRA_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                     606 rHisHisHisHisHis 612
                                                                                                                                                                                                                                                                                                                                                                                37 ACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Genomics 38:331-339(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MNB/DYRK SUBFAMILY.
                                                                                                                      56.50
4.708
                                                                                                                                                            52.174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mus musculus (Mouse)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  seq_documentation_block:
ID DYRA_MOUSE STAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=10090;
                                                                                                                                                              Percent Similarity:
                                                                                                                                             Ratio:
                                                                                                                          Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DYRK.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUBSTRATES
                                                                                                          alignment_scores:
                                                                                                                                                                                                   alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DYRK1A OR
  CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    region'
```

```
HSSP; P27703; LEKN.

MCD; MCI 1330299; Dyrkla.

InterPro; IPR000719; Euk_pkinase.

InterPro; IPR000729; Ser_thr_pkinase.

InterPro; IPR00059; Pkinase.

SMART; SM00220; S-TKC; 1.

PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE; PS00101; PROTEIN_KINASE_LOM; 1.

Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase; Arp-binding; Nuclear protein; Phosphorylation.

Arp-binding; Nuclear protein; Phosphorylation.

Arp-binding; Nuclear protein; Phosphorylation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostoml;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-CT-201 (Protein kinase minibrain homolog) (MNBH) (RP86) (Dual specificity YAK1-related kinase).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STRAIN-SPRAGUE-DANLEY: TISSUE-Brain;
MEDLINE-96216443; PubMace-8631952;
Kentrup H., Becker W., Heukelbach J., Wilmes A., Schuermann A.,
Huppertz C., Kainulainen H., Joost H.-G.;
"Dyrk, a dual specificity protein kinase with unique structural
features whose activity is dependent on tyrosine residues between
subdomains VII and VIII.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         590 HisValAlaProGlnGlnAsnAlaLeuHisHisHisHisGlyAsnSerSe 606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               60 CATGTCTTACCAAAA......AATAAGCT 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        E117DDD6C5E8C74F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps: 1
Percent Identity: 43.478
                                                                                                                                                                                                                                                                                                                          PROTEIN KINASE.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A., AND MUTAGENESIS OF TYR-219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            763 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Align seg 1/1 to: DYRA_MOUSE from: 1 to: 763
                                                                                                                                                                                                                                                                                                                                                                                                 BY SIMILARITY.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SER/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-HIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POLY-HIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             85494 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US-09-674-779-4/rev x DYRA_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          606 rHisHisHisHisHis 612
EMBL; U58497; AAC52994.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 37 ACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_name: SwissProt_40:DYRA_RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4.708 52.174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                  173
188
287
                                                                                                                                                                                                                                                                                                                                                                                                                       515
602
619
672
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             763 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DYRK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Ratio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DYRK1A OR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DYRA_RAT
063470;
                                                                                                                                                                                                                                                                                                                                                                                                    ACT_SITE
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOD_RES
SEQUENCE
                                                                                                                                                                                                                                                                                                                               DOMAIN
NP_BIND
BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MOD_RES
MOD_RES
```

|||||| |||::: 590 HisValAlaProGlnGlnAsnAlaLeuHisHisHisHisGlyAsnSerSe 606 60 CATGICITACCAAAA......AAIAAGCI 38

seq\_name: SwissProt\_40:PC\_DROME

seq\_documentation\_block:

PC DROME

to: 763

Align seg 1/1 to: DYRA\_RAT from: 1

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation between the Broppen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HSSP; P4//US; LEAR.
InterPro; IPR000719; Euk_pkinase.
InterPro; IPR000729; Ser_thr_pkinase.
Pfam; PF00069; pkinase; Z.
SMARY; SMO0220; S. TKE: 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00101; PROTEIN_KINASE_DOM; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
ATanaferase; Serine_threonine-protein kinase; Tyrosine-protein kinase; ATP-binding; Nuclear protein; Phosphorylation; Alternative splicing.
DOMAIN 117 134 ...
**COMPARTITE NUCLEAR LOCALIZATION SIGNAL ...
**COMPARTITE NUCLEAR ...
*
                                                                                                                                                                                                                                       "Sequence characteristics, subcellular localization, and substrate specificity of DYRK-related kinases, a novel family of dual specificity protein Kinases."

J. Biol. Chem. 273:25893-25902(1998).

- FUNCTION: MAY PLAY A ROLDE IN A SIGNALING PATHWAY REGULATING NUCLEAR FUNCTIONS OF CELL PROLIFERARATION. PHOSPHORYLATES SERINES, THREONINES AND TYROSINES RESIDUES IN ITS SEQUENCE AND IN EXOGENOUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TISSUE SPECIFICITY: UBIQUITOUS.
PTM: AUTOPHOSPHORYLATED ON TYR RESIDUES (BY SIMILARITY).
SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PHOSPHORYLATION (AUTO-).
MISSING (IN SHORT ISOFORM).
Y->F: REDUCED AUTOPHOSPHORYLATION ON
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TYROSINE, BUT NO LOSS OF HISTONE PHOSPHORYLATION.
CB5EC7EC4C1F9A47 CRC64;
                                                                                                                                                                                                Becker W., Weber Y., Wetzel K., Eirmbter K., Tejedor F.J., Joost H.-G.;
                                                                        Kentrup H.;
Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PHOSPHORYLATION (AUTO-).
PHOSPHORYLATION (AUTO-).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEIN KINASE.
ATP (BY SIMILARITY).
ATP (BY SIMILARITY).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   BY SIMILARITY.
POLY-SER.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SER/THR-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLY-HIS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLY-HIS
  Biol. Chem. 271:3488-3495(1996)
                                                                                                                                                                                                                                                                                                                                                                                                                                               SUBCELLULAR LOCATION: Nuclear. ALTERNATIVE PRODUCTS: 2 ISOFOR
                                                                                                                                                                      MEDLINE=98421512; PubMed=9748265;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           85541 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EMBL; X79769; CAA56164.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MNB/DYRK SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               219
319
321
78
219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        763 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P27703; 1ERK
                                                                                                                                                 CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1159
1188
1188
12887
5599
607
605
656
656
219
321
70
                                                                                                                                                                                                                                                                                                                                                                                                                             SUBSTRATES
                                                   REVISIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BINDING
ACT_SITE
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        VARSPLIC
MUTAGEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NP_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RES_
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Ş
```

Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

P26017; Q9VP49; 01-MAY-1992 (Rel. 22, Created) 01-MAY-1992 (Rel. 22, Last sequence update) 11-MAY-1992 (Rel. 22, Last sequence update) Polycomb protein. PC OR CG7618. Drosophila melanogaster (Fruit fly).

PRT;

STANDARD;

Paro K., Hogness D.S.;
"The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila.";
Proc. Natl. Acad. Sci. U.S.A. 88:263-267(1991).

MEDLINE=91095442; PubMed=1898775;

SEQUENCE FROM N.A.

STRAIN=OREGON-R;

NCBI\_TaxID=7227;

, Hogness D.S

Paro R.

```
RA Mannin-Zuujoudo; Pubade-10731132;

RA Adams M.D., Cellniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
Gocaye R.A., Lewis S.E., Richadel M.D., Zahang O., Chen L.X.,
Sutton G.G., Wortman J.R., Yandell M.D., Zahang O., Chen L.X.,
Randon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Bandon R.C., Baxter E.G., Helt G., Nelson C.R., Milkos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Barlew R.M., Basu A., Baxendale J., Bayratarogul L., Basaley E.M.,
RA Burkova D., Botchan M.R., Bouck J., Bayratarogul L., Basaley E.M.,
RA Gherry J.W., Cawley P.V., Berman B.P., Bhandari D., Botcher P.,
RA Cherry J.W., Cawley S., Dalkke C., Davenport L.B., Davies P.
RA Cherry J.W., Cawley S., Dalkke C., Davenport L.B., Davies P.
RA Cherry J.W., Cawley S., Dalkke C., Davenport L.B., Davies P.
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriara S., Fleischman W.,
RA Gorbert W. Doup L.E., Downes M., Dugan-Rocha S., Placks J.,
RA Goldek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harrey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Harrey D., Heiman T.J., Hernandez J.R., Elusky D. Lai Z.,
Alali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Alalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Alalush F., Milshina N.Y., Morris J., Mosher D. L.,
RA Mannel B.E., Kodire G.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Alalazolo M., Pittman G.S., Pan S., Pollad J., Worles D. M.,
RA Ralazzolo M., Pittman G.S., Pan S., Pollad J., Wang A.,
Ra Ralazzolo M., Pittman G.S., Pan S., Pollad J., Wang R.,
Spier E., Spradling A.C., Stapleton M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong S., Yao Q.A.,
Walliams S.M., Woodage T., Worley G., Wissesnbach J.,
Walliams S.M., Woodage T., Wolley G., Wang S., Yao Q.A.,
Rheng X.H., Zhong F.N., Zhong W., Zhu S., Zhu X., Zhu X., Zhu X., Zhu X., Zhu X., Zhu Z., Zh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-20196006; Pubmed=10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRAIN-BERKELEY
```

vaps: 1 Percent Identity: 43.478

56.50 4.708 52.174

Ratio:

Quality:

alignment\_scores:

Percent Similarity:

alignment\_block: US-09-674-779-4/rev x DYRA\_RAT

Length:

```
Developmental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FXGB_HUMAN P55315;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   alignment_block
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA_BIND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61 0 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1 T 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A D
    This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                 between the Swiss Institute of Bioinformatics and the EMBL outstation—the Buropean Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

"The genome sequence of Drosophila melanogaster.";

Science 287:2185-2185(2000).

-!- FUNCTION: REQUIRED FOR THE SEGMENT-SPECIFIC REPRESSION OF HOMEOTIC

SELECTOR GENES. PC MAY BE INVOLVED IN THE STABLE TRANSMISSION OF A DETERMINED STATE BY ITS EFFECTS ON CHROMATIN STRUCTURE. PROMOTES

LOCUS-SPECIFIC CHROMATIN COMPACTION.

-!- DEVELOPMENTAL STAGE: REQUIRED DURING THE ENTIRE LARVAL PERIOD FOR NORMAL ADULT DEVELOPMENT. IT IS FOUND IN ALMOST ALL CELLS AND TISSUES THROUGHOUT GASTRULATION AND ORGANOGENESIS THOUGH AT A MUCH LOMER LEVEL THAN IN BARLY SYMOYTIAL STAGES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TO MODIFIER 3 PROTEIN (AA 488-517).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-0CT-1996 (Rel. 34, Created)
01-0CT-1996 (Rel. 34, Last sequence update)
116-0CT-2001 (Rel. 40, Last annotation update)
Forkhead box protein G1A (Forkhead-related protein FKHL2)
(Transcription factor BF-2) (Brain factor 2) (BF2) (HFK2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5DB24AE4B326C3B9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Percent Identity: 80.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    469 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Align seg 1/1 to: PC_DROME from: 1 to: 390
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-HIS. POLY-HIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         130 LysGluLysLysHisHisHisHisHis 139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         48 AAAAATAAGCTACACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CHROMO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRINTS; PR00504; CHROMODOMAIN. SMART; SM00298; CHROMO; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PROSITE; PS00598; CHROMO_1; 1. PROSITE; PS50013; CHROMO_2; 1.
                                                                                                                                                                                                                                                                                                                         EMBL; X55702; CAA39229.1; -.
EMBL; AE003594; AAF51707.1; -.
PIR; A38565; A38565.
HSSP; P23197; 1AP0.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         43976 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       seq_name: SwissProt_40:FXGA_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                      FlyBase; FBgn0003042; Pc.
InterPro; IPR000953; Chromo.
Pfam; PF00385; chromo; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-674-779-4/rev x PC_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6.222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      56.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          143
167
380
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          134 1
160 1
351 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                         TRANSFAC; T00693;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FOXG1A OR FKHL2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TISSUE=Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                FXGA_HUMAN
P55316;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SIMILAR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
```

SOTETT TAND DRADE BY THE TENT TO THE SOLUTION OF THE SOLUTION

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                           MEDLINE-9532450; PubMed=7599184; MEDRINE-9532450; PubMed=7599184; MEDLINE-9532450; PubMed=7599184; Mariese S., Murphy D.B., Schlung A., Burfeind P., Schmundt D., Schnulle V., Mattei M.-G., Thies U.; and 2, members of the fork head gene family, are clustered on chromosome 14q."; are clustered on chromosome 14q."; elochim. Biophys. Acta 1262:105-112(1995).

-!-FUNCTION PLAYS AN IMPORTANT ROLE IN THE ESTABLISHMENT OF THE REGIONAL SUBDIVISION OF THE DEVELOPING BRAIN AND IN THE DEVELOPING BRAIN AND IN THE DEVELOPING SPECIFIC DNA-BINDING PROTEIN WITH A DISTINCT BINDING SPECIFICITY (BY SIMILARITY).
-!-SUBCELLULAR LOCATION: NUClear (Potential).
MEDLINE-9504833; PubMed-7959731; Murphy D.B., Wiese S., Burfeind P., Schmundt D., Mattei M.-G., Schulz-Schaeffer W., Thies U.; "Human brain factor 1, a new member of the fork head gene family.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              45 ProGlnHisHisHisHisHisHisHisHisHisHisProProPr 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51 CCAAAAAATAAGCTACACCATCACCATCACCATTAAGTCGACCCTCTGCC 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Transcription regulation; DNA-binding; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          943B8BDB90008EDC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Identity: 52.941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  477 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Align seg 1/1 to: FXGA_HUMAN from: 1 to: 469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-GLN.
POLY-ARG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FORK-HEAD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PROSITE: PS00657; FORK_HEAD_1; 1. PROSITE: PS00658; FORK_HEAD_2; 1. PROSITE; PS50039; FORK_HEAD_3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X78202; CAA55038.1; --
EMBL; X74143; CAA52240.1; --
HSSP; Q63245; ZHFH.
TRNSPCAC; T02292; --
MIM: 600779; --
InterPro; IPR001766; Fork_head.
PRIMTS; PR000559; FORK_HEAD.
SMART; SM00339; FH; I.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          50539 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-674-779-4/rev x FXGA_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    seq_name: SwissProt_40:FXGB_HUMAN
                                                                                                                     Genomics 21:551-557(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Quality: 56.00
Ratio: 5.091
Percent Similarity: 64.706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   56
79
73
88
252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   33
57
70
83
161
269 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          seq_documentation_block:
```

51 CCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGACCTCTGCC

N

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                              MEDLINE=5504313; PubMed=7959731; Murphy D.B., Wattei M.-G., Schulz-Schaeffer W., Thies U.; "Human brain factor 1, a new member of the fork head gene family."; Genomics 21:551-557(1994).
                                                                            Homo sapiens (Human).
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
01-0CT-1996 (Rel. 34, Created)
1-0CT-1996 (Rel. 34, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Forkhead box protein G1B (Forkhead-related protein FKHL1)
(Transcription factor BF-1) (Brain factor 1) (BF1) (HFK1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE; PS00657; FORK HEAD_1; 1.
PROSITE; PS00658; FORK_HEAD_2; 1.
PROSITE; PS50039; FORK_HEAD_3; 1.
Transcription regulation; DNA-binding; Nuclear protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          71CFD0BD069CFAD5 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLY-GLN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                POLY-LEU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              HIS-RICH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRO-RICH
                                                                                                                                                                                                                                                                      MEDLINE=95322450; PubMed=7599184;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           InterPro; IPR001766; Fork_head.
Pfam; PF00250; Fork_head; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             259 F
51340 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; X74142; CAA52239.1; -. HSSP; Q63245; 2HFH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRINTS; PR00053; FORKHEAD.
SMART; SM00339; FH; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSFAC; T02350; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            168 ;
477 AA;
                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                 FOXGIB OR FKHL1.
                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Developmental
                                                                                                                                                        TISSUE=Brain;
                                                                                                                                                                                                                                                            TISSUE-Brain;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MIM; 164874
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
```

SEQUENCE FROM N.A.
MEDLINE=91232862; PubMed=2030914;
Blake T.J., Shapiro M., Morse H.C. III, Langdon W.Y.;
"The sequences of the human and mouse c-cbl proto-oncogenes show
v-cbl was generated by a large truncation encompassing a proline-rich domain and a leucine zipper-like motif.";

Oncogene 6:653-657(1991)

Eukaryota; Metazoa; Chordata; Cranlata; Vertebrata; Euteleostomi; Mammalla; Eutheria; Primates; Catarrhini; Hominidae; Homo.

(Human).

Homo sapiens

NCBI\_TaxID=9606;

01-AUG-1991 (Rel. 19, Created)
01-AUG-1991 (Rel. 19, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Signal transduction protein CBL (Proto-oncogene c-CBL).
CBL OR CBL2.

PRT;

STANDARD;

seq\_name: SwissProt\_40:CBL\_HUMAN

seq\_documentation\_block:

CBL\_HUMAN P22681;

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -i- FUNCTION: PARTICIPATES IN SIGNAL TRANSDUCTION IN HEMATOPOIETIC CELLS. ADAPTOR PROTEIN THAT FUNCTIONS AS A NEGATIVE REGULATOR OF MANY SIGNALLING PATHWAYS THAT START FROM RECEPTORS AT THE CELL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MEDLINE=99176421; PubMed=10078535;
Meng W., Sawasdikosol S., Burakoff S.J., Eck M.J.;
"Structure of the amino-terminal domain of Cbl complexed to its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SMART; SN00184; RING; 1.
SMART; SM00165; UBA; 1.
PROSITE; PS00518; ZF_RING_1; 1.
PROSITE; PS50089; ZF_RING_2; 1.
3D-roto-oncogene; Nuclear protein; Zinc-finger; Phosphorylation; 3D-structure.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -!- SUBUNIT: ASSOCIATES WITH NCK VIA ITS SH3 DOMAIN.
-!- SUBCELLULAR LOCATION: Nuclear.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -!- PTM: PHOSPHORYLATED ON TYKUSINE.
-!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
-!- SIMILARITY: CONTAINS 1 UBA DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   X-ŘAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 47-350.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SUBCELLULAR LOCATION: Nuclear. PIM: PHOSPHORYLATED ON TYROSINE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      binding site on ZAP-70 kinase.";
Nature 398:84-90(1999).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR003153; Cbl_N.
InterPro; IPR000449; UBA.
InterPro; IPR001841; Znf_ring.
Pfam; PF02262; Cbl_N; 1.
Pfam; PF02761; Cbl_N2; 1.
Pfam; PF02761; Cbl_N3; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X57110; CAA40393.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PF00627; UBA; 1.
PF00097; zf-C3HC4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A43817; A43817.
1B47; 27-APR-99.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SURFACE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A43817;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MIM; 165360;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PIR;
A WAR REAL PROPERTY OF THE PRO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Identity: 52.941
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Align seg 1/1 to: FXGB_HUMAN from: 1 to: 477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-674-779-4/rev x FXGB_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5.091
64.706
```

Quality:

alignment\_scores:

Ratio:

Percent Similarity:

alignment\_block:

SHEEFFE

```
KKW
FT
FT
FT
FT
FT
SQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics institute. There are no restrictions on its muse by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NUCLEAR LOCALIZATION SIGNAL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Voisard C.P., Wang J., McEvoy J.L., Xu P., Leong S.A.;
"urbs1, a gene regulating siderophore biosynthesis in Ustilago
maydis, encodes a protein similar to the erythroid transcription
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Eukaryota, Fungi, Basidiomycota, Ustilaginomycetes;
Ustilaginomycetidae, Ustilaginales, Ustilaginaceae, Ustilago.
                                                                                                    PHOSPHORYLATION.
7D686B050204AD8F CRC64;
                                                                                                                                                                                  Gaps: 0
Percent Identity: 66.667
                              ASP/GLU-RICH (ACIDIC).
                                                          ASP/GLU-RICH (ACIDIC).
                                                                                                                                                                         12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Siderophore biosynthesis regulatory protein URBS1.
                                                                                                                                                                                                                                                                                                                                                                                                                                 01-FEB-1995 (Rel. 31, Created)
01-FEB-1995 (Rel. 31, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
                                                                                       PHOSPHORYLATION
                                                                                                                                                                                                                                                                                                                                                                                                    950 AA
                                                                                                                                                                         Length:
                                                                                                                                                                                                                                                                                                                      36 HisHisHisHisHisHisHeuSerProHisPro 47
                                                                                                                                                                                                                                                                                                    36 CACCATCACCATCACCATTAAGTCGACCCTCTGCCT 1
                                                                                                                                                                                                                                                                         Align seg 1/1 to: CBL_HUMAN from: 1 to: 906
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRINTS; PRO0619; GATAZNFINGER.
SMART; SMO0401; ZnF_GATA; 2.
PROSITE; PSO0344; GATA_ZN_FINGER_1; 2.
PROSITE; PS50114; GATA_ZN_FINGER_2; 2.
                                              PRO-RICH
                                                                                                                                                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=94019380; PubMed=8413298;
                                                                     UBA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HSSF, P17679; 1GNF.
TRANSFAC; T02406; -.
InterPro; IPR000679; ZnF_GATA.
InterPro; IPR001164; Znf_GCS.
                                                                                                                   MM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ustilago maydis (Smut fungus)
                                                                                                                                                                                                                                                                                                                                                            seq_name: SwissProt_40:URB1_USTMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EMBL; M80547; AAB05617.1; -.
                                                                                                                                                                                                                                              US-09-674-779-4/rev x CBL_HUMAN
                                                                                                                  99646
                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                    83.333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00320; GATA; 2
                                                                                                                                                                                        5.600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PIR; S27473; S27473.
                                                                                                                                                                                                                                                                                                                                                                                           seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                  906 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=5270;
                                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                        Ratio:
                                                         689
856
700
774
                                                                                                                                                                           Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                01-FEB-1995
16-OCT-2001
                                                                                                                                                             alignment_scores
                                                                                                                                                                                                                                  aliqnment_block:
                                                                                                                                                                                                                                                                                                                                                                                                         URB1_USTMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=518
                                                                                                                     SEQUENCE
                                                                                     MOD_RES
MOD_RES
   DOMAIN
ZN_FING
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                     P40349:
                                                          DOMAIN
DOMAIN
                                              DOMAIN
```

HID DE REAL STORY OF THE PROPERTY OF THE PROPE

```
COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORESORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY (BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
TISSUE-Electric lobe;
MEDLINE-93248175; PubMed=7683405;
Medlevine W.A., Ellino P.T., Inman I., Zhou M., Tsien R.W., Schwarz T.L.;
"Molecular diversity of Ca2+ channel alpha 1 subunits from the marine ray Discopyge ommata.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -1- TISSUE SPECIFICITY: EXPRESSION IS HIGHER IN THE ELECTRIC LOBE THAN IN THE POREBRAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS SPORDABLY. REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION. PTM: PHOSPHORYLATED IN VITRO BY CAM-KINASE II, CAPK, PKC AND CGPK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -i- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
-i- ALTERNATIVE PRODUCIS: AT LEAST 2 ISOFORMS ARE PRODUCED BY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Discopyge ommata (Electric ray).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
Elasmobranchii; Squalea; Hypnosqualea; Pristiorajea; Batoidea;
Torpediniformes; Narcinoidei; Narcinidae; Discopyge.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-JUL-1999 (Rel. 39, Last sequence update)
30-MAY-2000 (Rel. 39, Last annotation update)
Probable voltage-dependent N-type calcium channel alpha-1B subunit
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Proc. Natl. Acad. Sci. U.S.A. 90:3787-3791(1993).
-!- FUNCTION: THE ISOFORM ALPHA-1B GIVES RISE TO N-TYPE CALCIUM CURRENTS. N-TYPE CALCIUM CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP (BY SIMILARITY).
-!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
DNA-binding; Zinc-finger; Transcription regulation; Repressor;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
                                                                                                                                                                                       80 HIS-RICH.
06 POLY-SER.
101427 MW; F969C6DA09A78C12 CRC64;
                                                                                                                                                                                                                                                                                                                                                       Length: 13
Gaps: 0
Percent Identity: 61.538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  750 ValvalProHisHisHisHisHisHisHisHisHis 762
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              57 GICITACCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 2326 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Align seg 1/1 to: URB1_USTMA from: 1 to: 950
                                                       GATA-TYPE 1.
GATA-TYPE 2.
                                                                                                                                     POLY-SER.
POLY-THR.
                                                                                                              POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15-JUL-1999 (Rel. 38, Created)
15-JUL-1999 (Rel. 38, Last sequ
30-MAY-2000 (Rel. 39, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     US-09-674-779-4/rev x URB1_USTMA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             seq_name: SwissProt_40:CCAB_DISOM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ALTERNATIVE SPLICING.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                    76.923
                                                                                                                                                                                                                                                                                                                                                             56.00
                                                       362
506
27
36
490
780
                                                                                                                                                                                                                                                                                                                                                                                          5.600
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (BY SIMILARITY)
                                                                                                                                                                                                                                                       950 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
                                  Nuclear protein
                                                       338
482
24
28
487
753
902
                                                                                                                                                                                                                                                                                                                                                                  Quality:
                                                                                                                                                                                                                                                                                                                                                                                             Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SIMILARITY)
                                                                                                                                                                                                                                                                                                                                           alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CCAB_DISOM
P56698;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              alignment_block:
                                                                                                                                                                                                                                                       SEQUENCE
                                                                                    FING
                                                       ZN_FING
                                                                                                              DOMAIN
                                                                                                                                           DOMAIN
                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                             DOMAIN
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Burpaen Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@lsb-sib.ch).
S-FROT entry 1s cot,

the Swiss Institute o.

Jean Bioinformatics Institutions as.

ied and this statement is not remov.

Jean and this statement is not remov.

Jean and this statement is not remov.

Jean requires a licensediab-sib.ch).

Well: 112522; NOT_ANNORATED.CS.

INTERPTO: IPRO0021); Cachannel. TrpL.

INTERPTO: IPRO0025; Channel. Pore_Ca_Na.

JR FRINTS: PRO0131; Cachannel. TrpL.

INTERPTO: IPRO0026; Cation_Chanlel. TrpL.

INTERPTO: IPRO0026; Cation_Chanlel.

INTERPTO: IPRO0026
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CYTOPLASMIC (POTENTIAL).

S1 OF REPEAT I (POTENTIAL).

S2 OF REPEAT I (POTENTIAL).

S2 OF REPEAT I (POTENTIAL).

S3 OF REPEAT I (POTENTIAL).

S3 OF REPEAT I (POTENTIAL).

S4 OF REPEAT I (POTENTIAL).

S5 OF REPEAT I (POTENTIAL).

S5 OF REPEAT I (POTENTIAL).

S5 OF REPEAT I (POTENTIAL).

S7 OF REPEAT II (POTENTIAL).

S2 OF REPEAT II (POTENTIAL).

S3 OF REPEAT II (POTENTIAL).

S3 OF REPEAT II (POTENTIAL).

S4 OF REPEAT II (POTENTIAL).

S5 OF REPEAT II (POTENTIAL).

S4 OF REPEAT II (POTENTIAL).

S5 OF REPEAT III (POTENTIAL).

S1 OF REPEAT III (POTENTIAL).

S2 OF REPEAT III (POTENTIAL).

S3 OF REPEAT III (POTENTIAL).

S1 OF REPEAT III (POTENTIAL).

S2 OF REPEAT III (POTENTIAL).

S3 OF REPEAT III (POTENTIAL).

S4 OF REPEAT III (POTENTIAL).

S5 OF REPEAT III (POTENTIAL).

S6 OF REPEAT III (POTENTIAL).

S7 OF REPEAT III (POTENTIAL).

S6 OF REPEAT III (POTENTIAL).

S7 OF REPEAT III (POTENTIAL).

S6 OF REPEAT III (POTENTIAL).

S7 OF REPEAT III (POTENTIAL).

S7 OF REPEAT III (POTENTIAL).

S8 OF REPEAT III (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CYTOPLASMIC (POTEWILL).
SIOF REPEAT IV (POTEWILL).
SIOF REPEAT IV (POTEWILL).
SZ OF REPEAT IV (POTEWILL).
SY OF REPEAT IV (POTEWILL).
SY OF REPEAT IV (POTEWILL).
SZ OF REPEAT IV (POTEWILL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
TRANSMEM
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                DOMAIN
```

POLY-GLN. POLY-HIS. BINDING TO THE BETA SUBUNIT (BY SIMILARITY). CALCIUM TON SELECTIVITY AND PERMEABILITY (BY SIMILARITY). CALCIUM ION SELECTIVITY AND PERMEABILITY (BY SIMILARITY). PHOSPHORYLATION (BY CAPK) (POTENTIAL). BY SIMILARITY. N-LINKED (GLCNAC) (POTENTIAL). N-LINKED (GLCNAC) (POTENTIAL). D -> DOGLGIIYEPEQKPEDIOSYY (IN ISOFORM). SIS MY. DESAPRABUBLISH SELECTIVED.	Length: 23 Gaps: 1 rcent Identity: 43.478	<pre>ig 1/1 to: CCAB_DISOM from: 1 to: 2326 CATGTCTTACCAAAAAAAAAGCTA</pre>	19 2045 OUSE	RD; PRT; 323 AA.  Created) Last sequence update) Last annotation update) MAF1 (Segmentation protein KR) (Kreisler).  hordata; Craniata; Vertebrata; Buteleostomi;  odentia; Sciurognathi; Muridae; Murinae; Mus.	THE CAUGH END AND BY BAIN.  FUNCTION: MAY PLAY AN EARLY ROLE IN AXIAL PATTERNING (HINDBRAIN FUNCTION: MAY PLAY AN EARLY ROLE IN AXIAL PATTERNING (HINDBRAIN ENCYTION: MAY PLAY AN EARLY ROLE IN AXIAL PATTERNING (HINDBRAIN SEGMENTATION).  FUNCTION: MAY PLAY AN EARLY ROLE IN AXIAL PATTERNING (HINDBRAIN SEGMENTATION).  SEGMENTATION: NOT ABUNDANT IN KIDNEY, GUT, LUNG, AND BRAIN. DEVELOPMENTAL STAGE: DETECTABLE AT 8.0 DPC (ONE SOMITE) AS A BAND IN THE CAUDAL HINDBRAIN, AND BY 8.5 DPC (SIX TO EIGHT SOMITES), WITH THE CAUDAL HINDBRAIN, AND BY 8.5 DPC (SIX TO EIGHT SOMITES), WITH THE RA/RS BOUNDARY AND A DIFFUSE CAUDAL EDGE COINCIDENT WITH THE RA/RS BOUNDARY AND A DIFFUSE CAUDAL EDGE COINCIDENT SIMILARITY: BELONGS TO THE BZIP FAMILY. MAY SUBFAMILY.  SMISS-FROT entry is copyright. It is produced through a collaboration een the Swiss Institute of Bioinformatics and the EMBL outstation -
1873 2046 388 306 653 1362 1716 1716 1745 178 406 AA: 2644	5.50 625 .174 cab_D	9 1/1 to: CCAB_DISOM f CATGTCTTACCAAAAATAAGCTA       ::    HisValMetProAspTyrSerLeu	.CACCATCACCATCACCAT 1	block: STANDARD; Rel: 34, Created) Rel: 34, Last sec Rel: 34, Last and Last and factor MAFI (Seq CMOUSE): REML. Recas: Chordata; Razcas: Chordata; heria; Rodentia; N.A.	THE TOTAL SON TOTAL SON THE RAY OF THE RAY O
1869 2040 371 306 653 11362 1650 1734 271 1734 271 1558 406	: ity: tio: ity: /rev	to: C CTTACCA  :::	.CACCATCACCATCACCA 	STA	mouse segmen for transcript 79:1025-1034 FUNCTION: MAY SEGMENTATION) SEGMENTATION) SECELLULAR L TISSUE SPECIF IN THE CAUDAL THE HIGH LEVE WITH THE R4/R WITH THE R4/R WITH THE R4/R WITH THE R5.
DOMAIN DOMAIN SITE SITE SITE SITE CA_BIND CARBOHYD CARBOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD CARSOHYD	gnment_screent Singnment_bl	O O	36 .CACCATCACCATCACCAT 19	documentation_block:	Cordes S.P., Barsh G.S.  "The mouse segmentation fazipper transcription fa Cell 79:1025-1034(1994)
	alic Per alic US:	Al	-bes	Seq. 11D DT TO CO	8 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4

```
out_format : pfs
OM of: US-09-674-779-4 to: SPTREMBL_19:*
```

Date: Jul 30, 2002 4:31 PM

About: Results were produced by the GenCore software, version 4.5,

Copyright (c) 1993-2000 Compugen Ltd.

```
Database sequences: 562222
Database length: 172994929
Search time (sec): 47.140000
                                                                                                                            Search information block;
Command line parameters:
                                                                                                                                                       Database: SPTREMBL_19:*
                                                                                                                                     Query: US-09-674-779-4
Query length: 60
```

102 | QOSUIT A CALLALOUPES CONTINUE | 102 | QOVUEL | drosophila melanogaste | 219 | 101746 caenorhabditis elegans | 318 | QOSUF | Sanopheles gambiae (afr | 604 | O4247 zenopus laevis (african | 638 | QOVX9 drosophila melanogaste | 1186 | QOVX9 drosophila melanogaste | 165 | O17909 caenorhabditis elegan | 1012 | QOY45 drosophila melanogaste | 1012 | QOY45 drosophila melanogaste | 102 | QOY45 | QOY45 drosophila melanogaste | 102 | QOY45 | QOY55 | QOY45 | QOY55 243 ! 09p2r2 homo sapiens (human). hd 243 ! 09n2g3 pan troglodytes (chimpan Strd Orig ZSCore ESCOre Len 64.00 168.26 0.0475 10 64.00 168.26 0.0475 10 62.00 115.3 0.0807 40 62.00 175.00 0.2934 12 62.00 170.90 0.2934 12 62.00 170.90 0.2934 12 62.00 159.56 0.3151 48 62.00 151.71 0.3310 70.50.00 157.11 0.4621 45 70.00 157.11 0.4621 45 70.00 157.11 0.4621 45 70.00 147.78 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 77.8 0.4018 14 7 0.5007 221 58.00 143.69 0.5026 58.00 143.69 0.5030 58.00 143.61 0.5031 57.00 160.94 0.6513 57.00 157.20 0.6670 57.00 154.30 0.6790 57.00 154.30 0.6790 57.00 154.30 0.6793 57.00 154.30 0.6793 57.00 154.30 0.6793 57.00 154.30 0.6996 56.00 149.62 0.6996 56.00 157.06 0.9829 56.00 154.70 0.9829 56.00 148.29 1 1 56.00 143.21 1 56.00 143.40 1 1 56.00 139.40 1 1 55.50 153.19 1 1 4 55.00 153.19 1 1 4 55.00 153.19 1 1 4 58.00 14 58.00 14 58.00 14 58.00 14 58.00 14 58.00 14 57.00 16 sp\_lant:095871 56
sp\_lanvertebrate:0970E1 56
sp\_lanvertebrate:007346 59\_lanvertebrate:0980E0 5p\_lanvertebrate:0950Y5 5p\_lanvertebrate:0950Y5 5p\_lanvertebrate:09773 5p\_lanvertebrate:09773 5p\_lanvertebrate:0970Y5 5p\_lanverte sp\_invertebrate:09vz60 sp\_invertebrate:095z58 sp\_invertebrate:09v745 -sp\_invertebrate:022156 sp\_invertebrate:017909 sp\_invertebrate:024746 sp\_plant:09FJC5 sp\_vertebrate:09YGC6 sp\_human:09HCK8 sp\_plant:094F76 sp\_invertebrate:09N3Z0 sp\_vertebrate:Q90888 -sp\_vertebrate:Q90370 -sp\_human:Q9NSV0 sp\_vertebrate:091863 sp\_plant:09sG87
sp\_invertebrate:09GPH1 sp\_vertebrate:09PW44 -sp\_vertebrate:09PW46 -sp\_vertebrate:09PUM5 -sp\_vertebrate:09PUM4 sp\_vertebrate:09PW45 sp\_vertebrate:09PW47 sp\_vertebrate:09PUM6 Sp\_vertebrate:09PUM3 sp\_human:Q9P2R2 sp\_mammal:Q9N2G3 sp\_plant:09SN33 sp\_human:096F26 sp\_plant:09LOH2 score\_list:

```
pongo pygmaeus (orang
phaseolus vulgaris (k
homo sapiens (human).
     gorilla gorilla (gori
                                                                                                                                                                                                                                                                                      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                    Bevan M., Mueller M.W., Muendlein A., Felber R., Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.; Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mueller M.W., Muendlein A., Felber R., Mewes H.W., Lemcke K.,
Mayer K.F.X.;
                       09n2g1
041122
09y5q3
09h1f1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EU Arabidopsis sequencing project;
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AL035526; CAB37467.1; -.
EMBL; AL161549; CAB7884.1; -.
InterPro; IPR001270; CLP_AB.
InterPro; IPR000862; RFC.
INTERPRO; PR0300; CLPPROTEASEA.
Hypothetical protein.
SEQUENCE 1057 AA: 117111 MW; 0095A59A4B9EF66D CRC64:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1057 AA; 1171111 MW; 0095A59A4B9EF66D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EU Arabidopsis sequencing project;
Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                               01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
HYPOTHETICAL 117.1 KDA PROTEIN.
F28A21.230 OR AT4G18820.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
YOR3320W FROM CHROMSOME XV.
BAG7 OR YOR3320W OR YOR134W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Percent Similarity: 64.00 Length: 11
Ratio: 5.818 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 72.727
                                                                                                                                                      PRT; 1057 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      928 ProArgAsnAsnValHisHisHisHisHis 938
                                                                                                                                                                                                                                                                    Arabidopsis thaliana (Mouse-ear cress).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   51 CCAAAAATAAGCTACACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Align seg 1/1 to: Q9SN33 from: 1 to: 1057
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Saccharomyces cerevisiae (Baker's yeast)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
153.19
153.19
151.55
150.86
150.86
55.00
55.00
55.00
55.00
                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-674-779-4/rev x Q9SN33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    seq_documentation_block:
ID Q12128 PRELIMINARY;
                                                                                                 seq_name: sp_plant:09SN33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      seq_name: sp_fungi:Q12128
                                                                                                                                   seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 alignment_scores:
sp_mammal:09N2G2
sp_mammal:09N2G1
                               sp_plant:041122
sp_human:09Y503
sp_human:09H1F1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q12128
Q12128;
                                                                                                                                                   AC AC DIT DIT DIT DIT DIT OS
```

```
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Enchoperygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                 SEQUENCE FROM N.A. Voss H., Benes V., Rechmann S., Teodoru C., Schwager C., Paces V.,
                                                                                                                                                               Vlcek C.,
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhou L., Boulianne G.L.; "Comparison of the neuralized genes of Drosophila virilis and
                                                                                                                                                           Wiemann S., Rechmann S., Benes V., Voss H., Schwager C., Vloe
Stegemann J., Zimmermann J., Erfle H., Paces V., Ansorge W.;
Submitted (NOV-1995) to the EMBL/GenBank/DDBJ databases.
                                                              SEQUENCE FROM N.A.

Benes V., Andrade M.A., Rechmann S., Teodoru C., Banrevi A., Sander C., Valencia A., Ansorge W., Voss H.;

Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  302 HisMetLeuProArgSerArgAlaLeuSerAspSerAsnAsnPheThrIl 318
                                                                                                                                                                                                                                                                                  Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                               Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                          SGD; SQUUSEBU; BAG7.
InterPro; IPROG198; RhGGAP.
SMART; SM0620; RhGGAP; 1.
SMART; SM00324; RhGGAP; 1.
SEQUENCE 409 AA; 46216 MW; 2B0BBB21CD948E27 CRC64;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces NCBL_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    318 eHisHisHisHisHisHisHisHisAlaLeuPheProSerPro 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Identity: 35.484
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 36 ......CACCATCACCATCACCATTAAGTCGACCCTCTGCCT 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     747 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   to: 409
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. MEDLINE-95095077; PubMed-8001814;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Drosophila virilis (Fruit fly).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Align seg 1/1 to: Q12128 from: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  60 CATGTCTTACCAAAAAATAAGCTA..
                                                                                                                                                                                                                                                                                                                                                                 EMBL; X94335; CAA64053.1; -.
EMBL; X90518; CAA62109.1; -.
EMBL; Z75042; CAA99333.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       01-NOV-1996 (TrEMBLrel. 01, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    seq_name: sp_invertebrate: Q24746
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alignment_block:
US-09-674-779-4/rev x Q12128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62.50
3.906
51.613
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       seq_documentation_block:
ID Q24746 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           melanogaster.";
Genome 37:840-847(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                     SGD; S0005660; BAG7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NEURALIZED PROTEIN.
                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                  SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NEUR OR NEU.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        alignment_scores:
                                                                                                                                                                                                                                                                  Ansorge W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      024746;
```

```
01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
GENOMIC DNA, CHROMOSOME 5, Pl CLONE: MNC6.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosida II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
Sequence features of the regions of 1.013,767 bp covered by sixteen
physically assigned Pl and TAC clones.";
EMBL; AB015476; BAB09735.1; -.
InterPro; IFR003666; Auxin_inducible.
Pfam; PF02519; Auxin_inducible.
SEQUENCE 122 AA; 14332 MW; 94E6445BDAA3FCD3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
                                                                                                                                                                     747 AA; 81982 MW; C8E72569D6FDCA11 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      59.00 Length: 14
5.364 Gaps: 1
78.571 Percent Identity: 64.286
                                                                                                                                                                                                                                        Length: 14
Gaps: 0
Percent Identity: 64.286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCAT 19
EMBL, U12593; AAB60619.1; --
EMBL, U12593; AAB60619.1; --
EMBL, U12591, AAB60619.1; --
EMBL, U12591, AAB60619.1; JOINED.
EMBL, U12592, AAB60619.1; JOINED.
FlyBase; FB900013132; Dvir\neur.
FlyBase; FB900013132; Dvir\neur.
FlyBase; FB90001841; Znf_ring.
Fample FP00097; Zf-C3HC4; 1.
SMART; SMO0184; RING; 1.
                                                                                                                                                                                                                                                                                                                                                                                                           60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                               16 HisAsnLeuGlnGlnGlnGlnGlnLeuHisHisHisHisHis 29
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        122 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Align seg 1/1 to: Q9FJC5 from: 1 to: 122
                                                                                                                                                                                                                                                                                                                                                                         Align seg 1/1 to: Q24746 from: 1 to: 747
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=COLUMBIA;
MEDLINE=99087489; Pubmed=9872454;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   alignment_block:
US-09-674-779-4/rev x Q9FJC5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_documentation_block:
ID Q9FJC5 PRELIMINARY;
                                                                                                                                                                                                                                                             5.167
85.714
                                                                                                                                                                                                                                                                                                                                      US-09-674-779-4/rev x Q24746
                                                                                                                                                                                                                                              62.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               seq_name: sp_mammal:095LG8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     seq_name: sp_plant:09FJC5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ratio:
                                                                                                                                                                                                                                                                                Percent Similarity:
                                                                                                                                                                                                                                                Quality:
                                                                                                                                                                                                                                                                 Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           alignment_scores:
                                                                                                                                                                                                                              alignment_scores:
                                                                                                                                                                                                                                                                                                                       alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Tabata S.
                                                                                                                                                                           SEQUENCE
```

```
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,
RA Arandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani B.T., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Beson K.W., Bavos B.V., Berman B.P., Bhandari D., Bolshkov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Erottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chadra I.,
RA Cherry J.W., Cawley S., Dahike C., Davenport L.B., Davies P.,
RA Dodson K., Doup L.E., Downes M., Dugah-Rocha S., Dunkov B.C., Dunn P.,
RA Dodson K., Doup L.E., Downes M., Dugah-Rocha S., Dunkov B.C., Dunn P.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Fosler C., Gabriellian A.E., Garg V., Hernandez J.R., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibbeywam C.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                             Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Macaca.
                                                                                                                                                                                                                                                                                        Muramatsu S.;
"Excessive hand-wringing in a MPTP-treated monkey.";
"Excessive hand-wringing in a MPTP-treated monkey.";
Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF295597; AAK97131.1;
SEQUENCE 486 AA; 52426 MW; 3471B61D90D92A7D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      362 ProLysLysGluHisHisHisHisHisHisHisSerGluSerPro 376
                                   01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           51 CCAAAAATAAGCTACACCATCACCATCACCATTAAGTCGACCCT 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Similarity: 80.000 Percent Identity: 60.000
      486 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 1128 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Align seg 1/1 to: Q95LG8 from: 1 to: 486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-MAY-2000 (TrEMBLrel. 13, Created)
  PRT;
                                                                                                       METHYL-CPG-BINDING PROTEIN 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                seq_name: sp_invertebrate:09vz60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  alignment_block:
US-09-674-779-4/rev x Q95LG8
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     59.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           seq_documentation_block:
                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                     NCBI_TaxID=9541;
                                                                                                                                                                                                                                                                                                                                                                                                                                          alignment_scores:
   Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CG2186 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN-BERKELEY
0951G8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .09VZ60;
```

```
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Laixo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Shier E.C., Siden-Kiamos I., Simpson M., Stkopski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Spier E., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T.W., Weinst K.C., Mu D., Yang S., Yao O.A.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster.";
EMBL, AE003484, ARA479671;
DR FlyBase; FBgn0030243; CG2186.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Plasmodium falciparum (isolate 3D7).
Eukaryota: Alveolata: Apicomplexa: Haemosporida; Plasmodium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Boschet C.J., Vaquero C.M.; "Plasmodium falciparum krueppel-like gene."; "blasmodium falciparum krueppel-like gene."; Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases. EMBL; AJ31837; CAC42230.1; "BDB5.4J31937; CAC42230.1; "BDB5.4J564 MM; BBD631DE99DEFCC8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps: 0
Percent Identity: 61.538
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Percent Identity: 56.250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         910 HisteuAlaProSerSerLysGlnHisHisHisHis 922
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCAC 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 1266 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Align seg 1/1 to: Q9VZ60 from: 1 to: 1128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Align seg 1/1 to: Q95258 from: 1 to: 1266
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               seq_name: sp_invertebrate: 095258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     59.00
5.364
84.615
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-674-779-4/rev x 09VZ60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-674-779-4/rev x 095258
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       59.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             75.000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    KRUEPPEL-LIKE PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4.917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=36329;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STRAIN=3D7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KRUEPPEL1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q95Z58;
```

```
ID DT DT OC OC OC
                                                                                                                                                                                                                                                                                                                                                                Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
Amphibia; Batrachia, Anura, Mesobatrachia, Pipoidea, Pipidae;
                                                                                                                                                                                                                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MEDITINE-99282101; PubMed-9620779;
Jones P.L., Vernaak D., Kass S.U.,
Landsberger N., Strubboulis, J., Wolffe A.P.;
"Methylated DNA and MeCP2 recruit histone deacetylase to repress
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases. EMBL, AF106951; AAD03736.1; --
EMBL, AF051768; AAD02551.1; --
InterPro: IPR001739; MBD.
Pfam; PF01429; MBD; 1.
SWART: SM00391; MBD; 1.
SWART: SM00391; MBD; 1.
506 LysaAsnAsnLysHisHisHisHisHisLeuLeuLysTyrIlePro 521
                                                                                                                                                                                                                                                                                                                               Strausberg R.;
Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
EMBL; BC011695; AAH11695.1;
                                                                                                                                                                                                                                                                                                                                                                                                     454 AA; 50269 MW; 11FC856EF3EEF24E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-WAY-1999 (TrEMBLrel. 10, Created)
01-WAY-1999 (TrEMBLrel. 10, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
                                                                                                                       01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
UNKNOWN (PROTEIN FOR IMAGE:3534143) (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps: 0
Percent Identity: 64.286
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               60 CAIGICITACCAAAAATAAGCIACACCATCACCAICACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         467 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Strouboulis J., Wolffe A.P.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Align seg 1/1 to: Q96F26 from: 1 to: 454
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Xenopus laevis (African clawed frog).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT;
                                                                                          PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  METHYL-CPG-BINDING PROTEIN MECP2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         transcription.";
Nat. Genet. 19:187-191(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     seq_name: sp_vertebrate:Q9YGC6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58.00
5.273
78.571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        seq_documentation_block:
ID Q9YGC6 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-09-674-779-4/rev x Q96F26
                                                                         seq_documentation_block:
ID Q96F26 PRELIMINARY;
                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
TISSUE=LUNG CARCINOMA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Xenopodinae; Xenopus.
                                        seq_name: sp_human:Q96F26
                                                                                                                                                                                                        Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Xenopus MeCP2.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Quality:
Ratio:
                                                                                                                                                                                                                                                          NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=8355;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         aliqnment_block:
                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                         NON_TER
                                                                                                             296F26;
```

```
SEQUENCE FROM N.A.

MEDLINE=20450683; PubMed=10997877;

MEDLINE=20450683; PubMed=10997877;

Magase T., Kikuno R., Nakayama M., Ohara O.;

Magase T., Kikuno R., Makayama M., Introsawa M., Ohara O.;

Prediction of the coding sequences of unidentified human genes.

XVIII. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro.";

DNA Res. 7:273-281(2000).

EMBL; AB046784; BAB13390.1; -.

InterPro: IPR001650; Helicase_C.

Frank, PR00271; helicase_C;

SMART; SM00490; HELICC: 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NON_TER 1 1 1 SEQUENCE 1417 AA; 161269 MW; 199854F63D8BC6AE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
N-TYPE CALCIUM CHANNEL ALPHA-1B CDB4 VARIANT.
                                                                                                                                                                                                                                                                                                                                                                                                                              01-WAR-2001 (TrEMBLrel. 16, Created)
01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Percent Identity: 64.286
                                                                                                                                                                                                                                           14
                                                          Percent Identity: 71.429
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                60 CATGTCTTACCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                     54 TTACCAAAAATAAGCTA.....CACCATCACCATTACAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRT; 2146 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps:
              Length:
                                                                                                                                                                                                                                                                                                                                                                                  PRT; 1417 AA
                                      Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Align seg 1/1 to: Q9HCK8 from: 1 to: 1417
                                                                                                                                                                         Align seg 1/1 to: Q9YGC6 from: 1 to: 467
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      KIAA1564 PROTEIN (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         seq_name: sp_vertebrate:09PUM3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5.273
78.571
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-674-779-4/rev x Q9HCK8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            58.00
4.833
85.714
                                                                                                                                US-09-674-779-4/rev x Q9YGC6
                                                                                                                                                                                                                                                                                                                                                               seq_documentation_block:
ID 09HCK8 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            58.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                   seq_name: sp_human:Q9HCK8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Percent Similarity:
                                      Ratio:
Percent Similarity:
                Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        alignment_scores:
alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAR-2001
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     alignment_block:
                                                                                                          alignment_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KIAA1564.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CACHA1B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             09PUM3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          29PUM3;
                                                                                                                                                                                                                                                                                                                                                                                                          Q9HCK8;
```

Ŋ

```
Ratio: 5.273
Percent Similarity: 91.667
                                                                                                                                                                                                                                                                                                                            CACHA1B.
                                                                                                                                                                                                                                                    09PW46:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CACHA1B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9PUM5;
                                                                                                                                                                                                                                                    Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gallus gallus (Chicken).
Eukaryota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                     Lu Q., Dunlap K., "Cloning and functional expression of novel N-type Ca(2+) channel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 "Cloning and Furt."
"Cloning and Furt."
"Cloning and Furt."
"A Biol. Chem. 0:0-(1999).
EMBL: AF173019, AMD51822.1;
InterPro; IPR000636; Cation_chan_non_lig.
InterPro; IPR0002111; Cat_channel_TrpL.
InterPro; IPR0018207; Ca_channel_TrpL.
InterPro; IPR001862; Channel_pore_Ca_Na.
InterPro; IPR001696; Na_channel.
                                                                                                                                                                                                                                                                                                 2146 AA; 244650 MW; 281BBECAC494AE49 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5E85C3AAF9493589 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY-2000 (TrEWBLrel. 13, Created)
01-MAY-2000 (TrEWBLrel. 13, Last sequence update)
01-DEC-2001 (TrEWBLrel. 19, Last annotation update)
N-TYPE CALCIUM CHANNEL ALPHA-1B CDB8 VARIANT.
                                                                                                                                                                                                                                                                                                                                                                              Gaps: 0
Percent Identity: 58.333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2036 ValProThrAspGlnValHisHisHisHisHisHis 2047
                                                                                                                                                                                                                                                                                                                                                                Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 TTACCAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT; 2157 AA.
                                                                                                                                               J. Biol. Chem. 274:34566-34575(1999).

BREAL; AFI73015; AAD51818.1;

InterPro; IPR000636; Cation_chan_non_lig.

InterPro; IPR002111; Cat_channel_TrpL.

InterPro; IPR002077; Ca_channel_TrpL.

InterPro; IPR001682; Channel_pore_Ca_Na.
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Align seg 1/1 to: Q9PUM3 from: 1 to: 2146
                                                      SEQUENCE FROM N.A.
TISSUE=DORSAL ROOT GANGLION NEURONS;
MEDLINE=20044702; PubMed=10574919;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
TISSUE-DORSAL ROOT GANGLION NEURONS;
Lu Q., Dunlap K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      245854 MW;
                                                                                                                                                                                                                                                     Pfam; PF00520; ion_trans; 4
PRINTS; PR00167; CACHANNEL.
PRINTS; PR00170; NACHANNEL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Pfam; PF00520; ion_trans; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRINTS; PR00167; CACHANNEL.
PRINTS; PR00170; NACHANNEL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        seq_name: sp_vertebrate:09PW44
                                                                                                                                                                                                                                                                                                                                                                           5.273 91.667
                                                                                                                                                                                                                                                                                                                                                                                                                                   US-09-674-779-4/rev x Q9PUM3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                           58.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2157 AA;
                               NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                           Quality:
Ratio:
                                                                                                                                                                                                                                                                                                                                                                                    Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          alignment_scores:
Quality:
                                                                                                                                                                                                                                                                                                                                           alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                   alignment_block:
                                                                                                                                 variants.
                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CACHA1B.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 09PW44;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    09PW44
```

Length:

58.00

```
Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Cloning and Functional Expression of Novel N-type Ca2+ Channel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Variants.";
J. Biol. Chem. 0:0-0(1999).

BMBL; AFIJ3017, AAD51820.1, ...

R InterPro; IPR002111; Cat_channel_trpL.

InterPro; IPR002111; Cat_channel_trpL.

InterPro; IPR001696; Na_channel_

R InterPro; IPR001696; Na_channel_

R InterPro; IPR001696; Na_channel_

R InterPro; IPR001696; Na_channel_

R PRIMTS; PR00170; NACHANNEL.

R PRIMTS; PR00170; NACHANNEL.

R PRIMTS; PR0170; NACHANNEL.

R PRIMTS; PR0170; NACHANNEL.
                                                                                                                                                                                                                                                                                                                                                                                      01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
N-TYPE CALCIUM CHANNEL ALPHA-1B CDB6 VARIANT.
Gaps: 0
Percent Identity: 58.333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-MAY-2000 (TrEMBLrel. 13, Last sequence update) 01-DEC-2001 (TrEMBLrel. 19, Last annotation update) N-TYPE CALCIUM CHANNEL ALPHA-1B CDB2 VARIANT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length: 12
Gaps: 0
Percent Identity: 58.333
                                                                                                                                                                                                          2047 ValProThrAspGlnValHisHisHisHisHisHis 2058
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        2061 ValProThrAspGlnValHisHisHisHisHisHis 2072
                                                                                                                                                                                                                                                                                                                                                   PRT; 2171 AA.
                                                                                                                                                                              54 TTACCAAAAAATAAGCTACACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54 TTACCAAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRT; 2182 AA.
                                                                                                                                 Align seg 1/1 to: Q9PW44 from: 1 to: 2157
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Align seg 1/1 to: Q9PW46 from: 1 to: 2171
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A. TISSUE=DORSAL ROOT GANGLION NEURONS;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last seq
01-DEC-2001 (TrEMBLrel. 19, Last anno
                                                                                                                                                                                                                                                                             seq_name: sp_vertebrate:09PW46
                                                           alignment_block:
US-09-674-779-4/rev x Q9PW44
                                                                                                                                                                                                                                                                                                                     seq_documentation_block:
ID Q9PW46 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alignment_block:
US-09-674-779-4/rev x Q9PW46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              seq_name: sp_vertebrate: Q9PUM5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5.273
91.667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      seq_documentation_block:
ID Q9PUM5 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     58.00
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Quality:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Percent Similarity:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ratio:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Lu Q., Dunlap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         alignment_scores:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ID DIT DOT OCC OCC OCC
```

```
L Q., Dunlap K.;

"Cloning and functional expression of novel N-type Ca(2+) channel

"Variants."

"Salol. Chem. 274:34566-34575(1999).

"Babl. AF1730147, AAD51817.1;

"R EMBL; AF1730147, AAD51817.1;

"R InterPro; IPR0001536; Cation chan non liq.

"InterPro; IPR002111; Cat_channel. TrpL.

"InterPro; IPR001695; Cation channel."

"InterPro; IPR001695; Cation channel."

"InterPro; IPR001695; Cation channel."

"InterPro; IPR001695; Na_channel."

"InterPro; IPR001695; Na_channel."

"InterPro; IPR00167; CACHANNEL."

"PRINTS; PR00170; NACHANNEL."

"PRINTS; PR00170; NACHANNEL."

"PRINTS; PR00170; NACHANNEL."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gallus gallus (Chicken).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
                                                                                                                                                                                     Lu Q., Dunlap K.; "Cloning and functional expression of novel N-type Ca(2+) channel
                                                                                                                                                                                                                                                      | validucs. | vali
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 262972 MW; E6FB221909BFD20B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            01-MAY-2000 (TrEMBLrel. 13, Created)
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
N-TYPE CALCIUM CHANNEL ALPHA-1B CDB3 VARIANT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Percent Identity: 58.333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2072 ValProThrAspGlnValHisHisHisHisHis 2083
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT; 2321 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  54 TIACCAAAAAATAAGCTACACCATCACCATCACCAT 19
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Align seg 1/1 to: Q9PUM5 from: 1 to: 2182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
TISSUE-DORSAL ROOT GANGLION NEURONS;
MEDLINE-20044702; Pubmed=10574919;
                                                                                       SEQUENCE FROM N.A.
TISSUE-DORSAL ROOT GANGLION NEURONS;
MEDLINE-20044702; PubMed=10574919;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        58.00
5.273
91.667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          seq_name: sp_vertebrate:Q9PUM4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       alignment_block:
US-09-674-779-4/rev x Q9PUM5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        seq_documentation_block:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2321 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Quality:
Ratio:
Percent Similarity:
Gallus.
NCBI_TaxID=9031;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     alignment_scores:
                                                                                                                                                                                                                                            variants.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CACHA1B.
```

Length:

58.00

alignment\_scores: Quality:

```
OM of: US-09-674-779-4 to: A_Geneseq_032802:* out_format : pfs
Date: Jul 30, 2002 4:32 PM
```

About: Results were produced by the GenCore software, version 4.5,

Copyright (c) 1993-2000 Compugen Ltd.

Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

WODEL-frame+n2p.model -DEV-xlp

-DB-A\_Geneseq\_032802 -QEMN-fastan -SUFFIX=01190.rag

-GAPOP-4.500 -GAPEXT-0.050 -MINMATCH-0.100 -LOOPCL-0.000

-LOOPEXT-0.000 -GAPOP-4.500 -GAPEXT-7.000 -YGAPOP-60.000

-XGAPEXT-60.000 -GAPOP-4.500 -GAPEXT-7.000 -YGAPOP-60.000

-XGAPEXT-60.000 -DELOP-6.000 -DELEXT-7.000 -START-1

-MATRIX-01190 -TRANS-human40.cdi -LIST=1000 -DCCALIGN-200

-YGAPCATC-0190 -TRANS-human40.cdi -LIST=1000 -DCCALIGN-200

-WATRIX-0190 -TRANS-human40.cdi -LIST=1000 -DCCALIGN-200

-WATRIX-0190 -TRANS-HIM-NIN-15 -ALIGN-15 -MODEL-LOCAL

-NORM-ext -HEAPSIZE=500 -MINLEN-0 -WAXLEN-200000000

-USER-SO9674779\_@CGN1\_1\_88 -NCPU-6 -ICPU-3 -LONGLOG

-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLEPXY -WAIT -THREADS=1

Search information block:
Query: US-09-674-779-4
Query length: 60
Database: A\_Ceneseq\_032802:\*
Database sequences: 747574
Database length: 111073796
Search time (sec): 53.240000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000 WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

score\_list: Sequence

Strd Orig ZScore EScore Len ! Documentation

•		
		•
		•

```
OM of: US-09-674-779-4 to: Issued_Patents_AA:* out_format : pfs
```

Date: Jul 30, 2002 4:32 PM

About: Results were produced by the GenCore software, version 4.5, Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

Search information block:
Query: US.09-674-779-4
Query: US.09-674-779-4
Query: Length: 60
Database: Issued\_Patents\_AA:\*
Database sequences: 231628
Database length: 24425594
Search time (sec): 22.730000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000 WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

Strd Orig

score\_list: Sequence

ZScore EScore Len ! Documentation

		·

```
OM of: US-09-674-779-4 to: PIR_71:* out_format : pfs

Date: Jul 30, 2002 4:33 PM

About: Results were produced by the GenCore software, version 4.5,

Copyright (c) 1993-2000 Compugen Ltd.

COPYIGHT (c) 1993-2000 Compugen Ltd.

COMMAN Line parameters:

-MODEL=frame+_n7p.model - DEV=xlp
-Q=Cqn2_1/UGSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/UGOSPTO_spcol/U
```

EScore Len ! Documentation

ZScore

Strd Orig

score\_list: Sequence

	·	

```
OM of: US-09-674-779-4 to: SwissProt_40:* out_format : pfs
```

Date: Jul 30, 2002 4:38 PM

About: Results were produced by the GenCore software, version 4.5, Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:
-MODEL=frame+-n2p.model -DEV-x1p
-Q=CQG702\_L/150FOC\_SPOO\_TOOG9074079/runat\_30072002\_151837\_8635/app\_query.fasta\_1.116
-Q=CQG702\_L/150FOC\_SPOO\_TOOG9074079/runat\_30072002\_151837\_8635/app\_query.fasta\_1.116
-DB=SwissProt\_40 -GFWT=fastan -SUFFIX=0.100 -LOOPEXT=0.000
-GAPEXT=0.050 -MINMATCH=0.100 -LOOPECL=0.000 -KGAPEXT=6.000
-GAPEXT=0.050 -GARPEXT=0.050 -KGAPOP=60.000 -KGAPEXT=60.000
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -YGAPEXT=60.000
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=60.000 -THR\_SCORE=quality
-THR\_NIN=15 -ALIGN=15 -NODE=LOCAL -OUTFMT=pfs -NORM=ext
-HEAPSIZE=500 -MINIEN=0 -MAXIEN=20000000000
-USER=USO9674779\_GCGN1\_1\_45 -NCPU=6 -ICPU=3 -LONGLOG
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THREADS=1

Search information block:
Query; US-09-674-779-4
Query length: 60
Database: SwissProt\_40:\*
Database sequences: 105224
Database length: 38719550
Search time (sec): 16.530000

WARN: XGAPOP and YGAPOP must be equal. Assuming YGAPOP=XGAPOP=60.000 WARN: XGAPEXT and YGAPEXT must be equal. Assuming YGAPEXT=XGAPEXT=60.000

ZScore EScore Len ! Documentation

score\_list: Sequence

Strd Orig

		,	
			•

	·			
			·	•
				•

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                        - protein search, using sw model
```

OM protein

July 30, 2002, 16:07:07; Search time 29.83 Seconds (without alignments) 930.890 Million cell updates/sec Run on:

US-09-674-779-2 Title: Perfect score:

1 MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250 Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

747574 seqs, 111073796 residues Searched:

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

1. /SIDSJ/gcgdata/nold-geneseqy-embl/AAlyB3.DAT:\*
4. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB3.DAT:\*
5. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB3.DAT:\*
6. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB3.DAT:\*
7. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB8.DAT:\*
8. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB8.DAT:\*
8. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB8.DAT:\*
9. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB8.DAT:\*
10. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
11. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
12. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
13. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
14. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
15. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
16. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
17. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
18. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
19. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
19. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
19. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
20. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
21. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
22. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
23. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
24. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
25. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
27. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
28. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\*
29. /SIDSJ/gcgdata/hold-geneseqy-embl/AAlyB9.DAT:\* /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:\* /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980. /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981. A\_Geneseq\_032802:\* 

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

		Description	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	Drosophila moland	Fechenichia meranog	F coli growth and	Drosophila molabor	Drosophila melanog	Drosophila malanog	Drosophila melanog	Neisseria meningi	Amino acid segment	Drosophila melanog
		ID	AAB60645		AAG98419	AAG99078	ABB71696	ABB66410	ABB67249	ABB67966	AAY74841	AAY37545	ABB70845
		DB	22	22	22	22	22	22	22	22	21	20	22
		Match Length DB	250	1377	320	320	1218	932	932	842	308	382	1596
*P	Query	Match	100.0	9.9	6.5	6.5	6.3	6.1	6.1	6.1	6.1	0.9	0.9
		Score	1355	89.5	88	88	83.5	83	83	82.5	82	81	81
	Result	No.	-	7	е	4	5	9	7	œ	6	10	11

New BASB120 polypeptides and polynucleotides from Moraxella catarrhalis strain American Type Culture Collection 43617, for use as therapeutic

WPI; 2001-159872/16. N-PSDB; AAF59797.

Human transferase Novel human diagno Neisseria meningit Arabidopsis thalia Oxalyl-Cob decarbo	Act. Y Posis Spension of the Posis Spension	Amino acid sequenc Amino acid sequenc Amino acid sequenc L. monocytogenes a Drosophila melanog Protein encoded by	Novel human diagno Novel human polypeptide Human protein SEQ Human protein SEQ Human Protein SEQ Human HSGTI protei Arabidopsis thalia Drosophila melanog	Cephalosporin anti S. avermitilis ave Streptomyces averm Streptomyces averm Topoisomerase II b Human cancer assoc Human CD40 recepto
				5
AAB73524 ABG12098 AAY74840 AAG1671 AAR75085 AAW22882	200000000	3 4 3 3 5 5 6 6	ABC07270 AAM41302 AAM41302 AAM39419 AAM39419 AAM39416 AAM3641672 ABB64072 ABB64072 ABB64072	AAK10145 AAB23749 AAG65264 AAG65268 AAY03181 AAB43576 AAY56019
755 755 757 757 757 757 757 757 757 757	22 22 23 23 23 23 23 23 23 23 23 23 23 2		200000000000000000000000000000000000000	71707777
233 304 388 578	307 342 348 2464 308	330 330 368 612 935 1783	1798 1139 1139	3972 3972 3972 3972 337 337
വസനസസസ യ്ങ്ങ്ങ്	7.		$\alpha$	ຸນທູດທູດທູດ ວັດດັດບັນບັ
80 78 78 78 78	77.5	· r r r r r r	9 9 9 9 9 9	75.5 75.5 75.5 75 75
12 13 14 15 17	222 222 322 332 332 332 332 332 332 332	222 222 24 25 26 26	33 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	044444 007 4443 207 207 207 207 207 207 207 207 207 207

## ALIGNMENTS

BASB120 protein; strain ATCC43617; antigen; antibody; vaccine; genetic immunisation; infection, upper respiratory tract; otitis media; hearing loss; deafness; pneumonia; sinusitis; nosocomial infection; invasive disease; antibacterial; auditory. Moraxella catarrhalis strain ATCC43617 BASB120 protein. (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS. AAB60645 standard; Protein; 250 AA. 31-JUL-2000; 2000WO-EP07361. 99GB-0018281 (first entry) Moraxella catarrhalis. WO200109335-A2 03-AUG-1999; 04-MAY-2001 Thonnard J; AAB60645; RESULT  ~

```
The invention learces to the Moraveira catalinaris strains in the intention also relates to immunogenic fragments of the BASB120 protein (AAB60645) and to DNA encoding it (AAF59797). The invention also relates to immunogenic fragments of the BASB120 protein, expression vectors and host cells comprising BASB120 nucleic acids, the recombinant production of BASB120, vaccine compositions comprising the compositions comprising the anti-BASB120 antibody, and a method of identifying a Moravella catarrhalis infection via the detection of BASB120 proteins or antibodies. The vaccine compositions of the invention are sprophylactic or therapeutic agents against Moravella catarrhalis is a Gram negative bacterium frequently isolated from the human upper respiratory tract, which is responsible for several catarchialis is a Gram negative bacterium frequently isolated from the human upper respiratory tract, which is responsible for several cases in children (which can lead to temporary or permanent hearing loss). It also causes pneumonia in elderly people, and sinusitis, no concomial infections and, less frequently, invasive diseases. BASB120 proteins or nucleotides may additionally be used in screening for novel and present sequence represents the Moravella catarrhalis strain aftering the Moravella catarrhalis strain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     121 EHLWGGIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDIWV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ehlwgqivptlhlyqdlksrgilpantqirsvyrnpelnqcaggaamskhltnsaldiwv 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PDLEIKSQALYELQNRLCQYWLEHGENQNFGLGLYATGAIHLDTQGFRKWGAQFSETNSI 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPP 120
agents or vaccines against bacterial infections, e.g. otitis media or
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1 MKNFNQYFITTLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHV 60
                                                                                                 to the Moraxella catarrhalis strain ATCC43617
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100.0%; Score 1355; DB 22; Length 250; 100.0%; Pred. No. 1.7e-145;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Drosophila melanogaster polypeptide SEQ ID NO 20910.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ABB64706 standard; Protein; 1377 AA
                                                           Claim 4; Page 64; 75pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26-MAR-2002 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Drosophila melanogaster
                                                                                                     invention relates
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             241 crhvlpknkl 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        241 CRHVLPKNKL 250
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      250 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            pharmaceutical.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches 250;
                          pneumonia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABB64706;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ABB64706
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
```

```
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                            New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -DLKSRGILPA 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       401 ttdwlmyhrkqnpyqvqpthcssttqssldsdasltpslgdfelksacsvdggskfgiga 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                292 piknssqpsplpkpktptikstkekarsldsaaneselsiv-----vhnitesh 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           79 -----STLGNYLPPMSQLLTTARSWQACG 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  146 NTQIRSVYRNPE------LNQC-----AGGAAMSKHLTNSAIDIWVPDLEIK 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      461 slaprsahkhngllhssstnlktlpecltlvefsssgggpkespfkgksmdlpmptlgak 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23 PIPTNPQVSPI---KIPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATKQ- 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Escherichia coli; identification; proliferation; microorganism; antimicrobial; antibacterial; antibiotic; gene therapy; diagnosis; bacterial growth inhibition.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (ABB57737-ABB72072).
The sequence data for this patent did not form part of the printed
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 1377;
                                                                                                                                                                                                                                                                                                                           Disclosure, SEQ ID NO 20910; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        113 ----HEPYQLPPEHLW------GQIVPTLHLYQ----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.6%; Score 89.5; DB 22;
18.8%; Pred. No. 2.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           87;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Escherichia coli protein sequence SEQ ID NO:467.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           42; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               187 SQALYELQNRLCQYWLEHGENQNFGLGLYAT 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    521 ttvstssmnll----grrgsnhsltlnlhss 547
                                                                                                                                                  Myers EW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAG98419 standard; Protein; 320 AA.
                                                                                                                                                  Li PWD,
                23-MAR-2001; 2001WO-US09231.
                                                  23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity 18.8%
Les 51; Conservative
                                                                                                                                                    Venter JC, Adams M,
                                                                                                                                                                                         WPI; 2001-656860/75.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1377 AA;
                                                                                                           (PEKE ) PE CORP NY
                                                                                                                                                                                                           N-PSDB; ABL08809
                                                                                                                                                                                                                                                                                        Interactions -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AAG98419;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAG98419
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            XX
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δχ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a
```

AAG99078 standard; Protein; 320 AA.

26-SEP-2001 (first entry)

```
The present invention describes a purified or isolated nucleic acid sequence (I) consisting essentially of one of the 93 nucleotide sequences cy given in AAH81202 to AAH81294, where expression of the nucleic acid in a microorganism is capable of inhibiting proliferation of a microorganism.

(I) have antibacterial and antibiotic activities, and can be used in comparism, and the manufactured antibiotic is useful for cof the microorganism, and the manufactured antibiotic is useful for reducing the activity or level of a gene product required for cof the microorganism in a subject, specifically humans. The reducing that inhibit bacterial growth or proliferation can be used as antisense therapeutics for killing bacteria. In addition to concern required for proliferation can be used as diagnostic tools.

(For example, nucleic acid probes complementary to proliferation required cor proliferation can be used as diagnostic tools.

(For example, nucleic acid probes complementary to proliferation-required cor be used as probes to identify particular species of microorganisms can colinical specimens. AAH81295 to AAH81489 to AAG98431, and AAH81489 to AAH81491.

(For example, nucleic acid probes complementary to proliferation corporations given in AAG98239 to AAG98431, and AAH81480 to AAH81491.
                                                                                                                                                                                                                                                                                                                                                             Novel nucleic acids encoding proteins required for Escherichia coliproliferation, useful for screening for antimicrobial agents -
                                                                                                                                                                                                                                                           Zyskind JW;
                                                                                                                                                                                                                                                                                                                                                                                                                             Example 4; Page 583; 596pp; English
                                                                                                                          19-DEC-2000; 2000WO-US34419.
                                                                                                                                                                      99US-0173005
                                                                                                                                                                                                                (ELIT-) ELITRA PHARM INC.
                                                                                                                                                                                                                                                      Forsyth RA, Ohlsen KL,
                                                                                                                                                                                                                                                                                                 2001-457376/49.
Escherichia coli.
                                                                                                                                                                                                                                                                                                                  N-PSDB; AAH81475
                                        WO200148209-A2.
                                                                                                                                                                  23-DEC-1999;
                                                                                    05-JUL-2001
```

68 TWLQMHHATKQEVVRYQAYLQSRLG--NYLPPMSQLLTTARSWQACGHEPYQLPPEHLWG 125 126 QIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCA------GGAAMSKHLTNSAI 176 90 --iplqtllpdapregiviniaelrlyyyppgknsvtvypigigqlggdtltptmvttvs 147 148 dkranptwtptanir--arykaggielpavvpagldnpmghhairlaayggvyllhgtna 205 84; Gaps 8 FITTLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67 55 ---sleaiakkynvgflallganpgvdpyvpragsvlt-----------Similarity 20.8%; Pred. No. 0.35; Similarity 20.8%; Pred. No. 0.35; Conservative 30; Mismatches 88; Indels 88; Indels 209 NFGLGL-YATGAIHL 222 :||:|: ::| | | 206 dfgigmrvssgcirl 220 Query Match Best Local Simi Matches 53; g δŏ 8 ŏ a ŏ οy

11;

320 AA;

Sequence

11;

88; Indels 84; Gaps

68 TWLQMHHATKQEVVRYQAYLQSRLG--NYLPPMSQLLTTARSWQACGHEPYQLPPEHLWG 125

126 QIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCA-----

AAG99078 RESULT

--::

8 FITTLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67

ŏ qq δŏ Q δy

10 wltfftfaaavalalpakantwplp-pagsrlvgenkf---hvvendgg--

---GGAAMSKHLTNSAI 176

AAA AAHB4373 to AAHB4499 represent Escherichia coli growth and proliferation related DNA sequences (I). AAHB4500 to AAHB4670 encode the E. coli related DNA sequences (I). AAHB4500 to AAHB4670 encode the E. coli growth and proliferation related proteins given in AAG99078 and AAG98030 to AAG98099. (I) can be used as potential targets for the generation of new antimicrobial agents, and for identification of compounds which interact with the gene products of (I). In addition the expression of the proteins, the purified proteins can be used to generate reagents and screen small molecule libraries or other candidate compound libraries for compounds that can be further developed to yield novel antimicrobial compounds. In addition, nucleic acid probes complementary to (I) that are specific for particular species of microorganisms can be used to identify particular microorganism species in clinical specimens, therefore, providing a rapid and dependable method by which to identify the causative agents of a bacterial method by which to identify the causative agents of a bacterial of method by which to identify the causative agents of a bacterial completion from proliferation-required sequences can also be used to screen for specific microorganisms that produce such proteins in a specimens as specific microorganisms that produce such proteins in a specimens used in the isolation of E. coli growth and proliferation in incortion in an example from the present E. coli growth and proliferation related protein sequence SEQ ID NO:299. Novel nucleic acids that inhibit Escherichia coli proliferation, useful for screening for homologous genes and for designing expression vectors Escherichia coli; growth; proliferation; microbial; antimicrobial; bacterial infection; microorganism. Query Match 6.5%; Score 88; DB 22; Length 320; Best Local Similarity 20.8%; Pred. No. 0.35; Matches 53; Conservative 30; Mismatches 88; Indels 6 Claim 19; Page 389-390; 522pp; English. 09-NOV-2000; 2000WO-US30950. 99US-0164415. (ELIT-) ELITRA PHARM INC. Forsyth RA, Ohlsen K, WPI; 2001-335933/35. N-PSDB; AAH84500. Sequence 320 AA; Escherichia coli. WO200134810-A2. 09-NOV-1999; 17-MAY-2001. 

q

δ

g à

```
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (ABB57737-ABB72072).
The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    90 RLGNYLPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQI 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       150 RSVY-----IMVPD 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pisrpiqmvs----eisesfdqisyqkgstvlrmmhlflgeesfrsglqaylqk 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ------fsyknaegdnlwesltgaahkyrs-----lpksydi 488
906 ---pvpnrllageivp---vhvtlrnmgiap----ieeiyl-----gcdnprslrn 946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    32 PIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATKQEVVR--YQAYLQS 89
                                                                                                                                                                                                                                                                                          Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Disclosure; SEQ ID NO 26022; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6.1%; Score 83; DB 22; Length 932; 20.0%; Pred. No. 6.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   65; Indels
                                                                                                                                                                                                                                                           Drosophila melanogaster polypeptide SEQ ID NO 26022.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred. No. 6.4;
; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Myers EW;
                                                                                                                                                            AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        20.08;
                                     174 SAIDIWVPDLËIKSQALYELQNR 196 : | | | | | | | | | | | | |
                                                                    947 lsndklvkdkeirggrvyrllnr 969
                                                                                                                                                        ABB66410 standard; Protein; 932
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Li PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                  23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                             (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                 Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Venter JC, Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        WPI; 2001-656860/75.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (PEKE ) PE CORP NY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   932 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N-PSDB; ABL10513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            interactions -
                                                                                                                                                                                                                                                                                                                   pharmaceutical
                                                                                                                                                                                                                             26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                  27-SEP-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                              ABB66410;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                          ABB66410
                                                                                                                          RESULT
                                                                                                                                                                                                                                               q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
       g
                                       QY
                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             . 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        856 dpaas11g-tlqfetqkirphnakqs---sqtvmdnrltiklvpq1pamnvsft----- 905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 EPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTN 173
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             54 DESVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGH 113
       --iplqtllpdapregiviniaelrlyyyppgknsvtvypigigqlggdtltptmvttvs 147
                                         --- RLCQY --- WLEHGENQ 208
                                                                       148 dkranptwtptanir--arykaqgielpavvpagldnpmghhairlaayggvyllhgtna 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                     Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match 6.2%; Score 83.5; DB 22; Length 1218; Best Local Similarity 27.3%; Pred. No. 8.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Disclosure; SEQ ID NO 41880; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           57; Indels
                                                                                                                                                                                                                                                                                                                                   Drosophila melanogaster polypeptide SEQ ID NO 41880.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    <del>..</del>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Myers EW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -
::
                                                                                                                                                                                                                                    ABB71696 standard; Protein; 1218 AA
                                            177 D----IWVPDLEIKSQALYELQN--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Li PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             23-MAR-2000; 2000US-191637P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          23-MAR-2001; 2001WO-US09231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11-JUL-2000; 2000US-0614150
                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                             209 NFGLGL-YATGAIHL 222
                                                                                                                                        206 dfgigmrvssgcirl 220
                                                                                                                                                                                                                                                                                                                                                                                                                         Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 2001-656860/75.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (ABB57737-ABB72072)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1218 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  N-PSDB; ABL15799
                                                                                                                                                                                                                                                                                                                                                                                                                                                        WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                       pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    interactions
                                                                                                                                                                                                                                                                                                    26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27-SEP-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 39;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Venter JC,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence
                                                                                                                                                                                                                                                                    ABB71696;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
            90
                                                                                                                                                                                                                    ABB71696
                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ٥y
```

.; 8

Gaps

```
26-MAR-2002
                                                                                                                                                                                                                                                                                27-SEP-2001.
                                                                                                                                                                                                                                                                                                                                                                                  Venter JC,
                                                                                                                                  ABB67966
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182
                                                                                                ABB67966
                                                                                       RESULT
         g
                               δy
                                                                                                                                 ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    .
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                               New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
489 ksimdswtlqtgypvinvtrdyaartaklng----eryllntqvarayrggcwwvp- 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             90 RLGNYLPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQI 149
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ------------------fsyknaeqdnlwesltgaahkyrs-----lpksydi 488
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----RNPELNQCAGGAAMSKHLTNSAID-----IWVPD 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         408 pisrpiqmvs-----eisesfdqisyqkgstv1rmmhlf1geesfrsg1qay1qk 457
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 76; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 32 PIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATKQEVVR--YQAYLQS 89
                                                                                                                                                                                      Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Disclosure; SEQ ID NO 28539; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       6.1%; Score 83; DB 22; Length 932; 20.0%; Pred. No. 6.4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                65; Indels
                                                                                                                                                                   Drosophila melanogaster polypeptide SEQ ID NO 28539.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ..
--
-:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               23; Mismatches
                                                                                                                                                                                                                                                                                                                                                                           Li PWD, Myers EW;
                      183 LEIKSQALYELQNRLCQYWLEHGEN 207
                                            541 lsyttqavqdfnntapkawmecgkn 565
                                                                                                 ABB67249 standard; Protein; 932 AA
                                                                                                                                                                                                                                                                                            23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                 23-MAR-2000; 2000US-191637P. 11-JUL-2000; 2000US-0614150.
                                                                                                                                              (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                          Drosophila melanogaster
                                                                                                                                                                                                                                                                                                                                                                         Adams M,
                                                                                                                                                                                                                                                                                                                                                                                              WPI; 2001-656860/75.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         150 RSVY-----
                                                                                                                                                                                                                                                                                                                                                   (PEKE ) PE CORP NY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (ABB57737-ABB72072)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Local Similarity
tes 41; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    932 AA;
                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; ABL11352
                                                                                                                                                                                                     pharmaceutical.
                                                                                                                                                                                                                                                WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                      interactions -
                                                                                                                                               26-MAR-2002
                                                                                                                                                                                                                                                                      27-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                           Venter JC,
                                                                                                                        ABB67249;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                        ABB67249
                                                                            RESULT
q
                                                                                                                      ŏ
                                           q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

```
capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and in the cell interactions in higher entaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL04840-ABL16175) and the encoded proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.w1po.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
489 ksimdswtlqtgypvinvtrdyaartaklng-----eryllntqvarayrggcwwvp- 540
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 86 YLQSRLGNYLPPMSQLLTTARSWQACGHEP-----YQLPPEHLW------GQ 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  126 pevlrnktksevlllgvasdrvrvphshsyppvgglghy----qqhhqqqqqqqqqqak 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         yaairrsyyaptqqq-----hnpqmpqhthaqqhmppthlmpppmmmmkssgs 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  127 IVPTLHLYQDLKSRGI-----LPANTQIRS---VYRNPELNQCAGGAAMSKHLT 172
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28 PQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATKQEVVRYQ--A 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   The invention relates to an isolated nucleic acid detection reagent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                :69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Disclosure; SEQ ID NO 30690; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 842;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Drosophila melanogaster polypeptide SEQ ID NO 30690.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 82.5; DE Pred. No. 6.3; 15; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Myers EW;
                                                                                                                                                                                                                                                                    ABB67966 standard; Protein; 842 AA
                                                           183 LEIKSQALYELQNRLCQYWLEHGEN 207
                                                                                                                       541 lsyttgavgdfnntapkawmecgkn 565
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Li PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           6.18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 21.8%;
                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    41; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Drosophila melanogaster.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ABB57737-ABB72072).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 2001-656860/75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (PEKE ) PE CORP NY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Sequence 842 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N-PSDB; ABL12069
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  interactions -
```

```
Wed Jul 31 08:19:56 2002
```

엄 ŏ

represent novel Neisseria meningitis and N. gonorrheae polynucleotides and polypeptides. AAZ54577 to AAZ54576 and AAZ54616 to AAZ5473 represent propertides. AAZ54577 to AAZ5476 and AAZ54616 to AAZ5473 represent propertides, antibodies and composition. The polypeptides, the polynucleotides, antibodies and compositions of the invention can be used as vaccines, as diagnostic reagents, and as immunogenic compositions. The polypeptides can be used in the manufacture of medicaments for treating or preventing infection due to Neisserial bacteria (e.g. meningitis and septicaemia), to detect the presence of Neisseria bacteria, or to raise antibodies. They may also be use as antibacterial agents. The polynucleotides of the invention may also be used in gene therapy protocols. Masignani V, Mora M; Scalato E, Scarselli M; Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine; antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia; antibacterial; gene therapy. 230 svppmeyrhsmmatgaptsymslnhmsspttkgamtssviyhgp-----pamshhts 281 Novel Neisserial polypeptides predicted to be useful antigens for Neisseria meningitidis ORF 280 protein sequence SEQ ID NO:1156. Hickey E, Ratti G, Claim 2; Page 638; 1453pp; English. Grandi G, | Rappuoli R, AAY74841 standard; Protein; 308 AA. 98US-0094869. 98US-0094869. 98US-0098994. 98US-0099062. 98US-0103749. 98US-0103794. 99WO-US09346. 98US-0103796. 99US-0121528 21-MAR-2000 (first entry) (GENO-) INST GENOMIC RES. vaccines and diagnostics Petersen J, Pizza M, l Pettelin H, Venter JC; Galeotti C, Neisseria meningitidis. WPI; 2000-062150/05. (CHIR ) CHIRON CORP 173 NSAIDIWV 180 282 lssstpwm 289 N-PSDB; AAZ53603 WO9957280-A2. 30-APR-1999; 19-OCT-1998; 25-FEB-1999; 01-MAY-1998 02-SEP-1998 32-SEP-1998 19-0CT-1998, 19-0CT-1998, Fraser C, Sequence AAY74841; AAY74841 RESULT 

```
AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome of Chlamydia trachomatis (see AAZ01425). The polypeptides can be used as vaccines against chlamydia trachomatis. Antisenses and ribozyme sequences can also be used to control growth of the microorganism. Chlamydia trachomatis is responsible for a large number of diseases, e.g. eye paratrachoma, and inclusion conjunctivitis; genital diseases such as nongonococcal uretritis, epidymitis, cervicitis, salpingitis, and venereal lymphogranulomatosis. The polypeptides of the invention may be of use in treating these diseases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Vaccine; eye disease; conventional trachoma; nonendemic trachoma; paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis; nongonococcal uretritis; epidymitis; cervicitis; salpingitis; bartholinitis; pneumopathy; venereal lymphogranulomatosis.
                             104 iqplk-----aeeegghhhdhdhdhdhdheg------hhhdhgeydphvwndpv 146
                                                                  75 ---ATKQEVVRY-----QAYLQSRLGNYLPPMSQL------LTTARSWQACGH 113
                                                                                                                                                                      207 dafsymgkryhiefiapqgysseaepsakqvaaiirgikregikavftenikdtrmvdri 266
                                                                                                                                     114 EPYQ-----LPPEHLWGQIVPTLH----LYQDLKSRGILPANTQ----IRSVYR- 154
                                                                                                   147 lmsayaqnvaealikadpegkvyyqqrlgnyqmqlkklhsdaqaafnavpaakrkvltgh
30 VSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHH--------
                                                                                                                                                                                                                                                                                                                                                                                                                                         Amino acid sequence of a Chlamydia trachomatis protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 1207-1209; 1755pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Genome sequence of Chlamydia trachomatis
                                                                                                                                                                                                               155 NPELNQCAGGAAMSKHLTNSAIDIWV 180
                                                                                                                                                                                                                                                  267 aketgvnvsgklysdalgnapadtyi 292
                                                                                                                                                                                                                                                                                                                                         AAY37545 standard; Protein; 382 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         98WO-IB01939.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           97FR-0015041.
97FR-0016034.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98US-0107077
                                                                                                                                                                                                                                                                                                                                                                                                             07-0CT-1999 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chlamydia trachomatis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 1999-371125/31.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (GEST ) GENSET
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO9928475-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            04-NOV-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                17-DEC-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Griffais R;
                                                                                                                                                                                                                                                                                                                                                                            AAY37545;
                                                                                                                                                                                                                                                                                                          RESULT 10
                                                                                                                                                                                                                                                                                                                          AAY37545
                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                    δŽ
                                                                           Qγ
                                                                                                           g
                                                                                                                                                δλ
```

11;

Gaps

84;

6.0%; Score 81; DB 20; Length 382; 18.9%; Pred. No. 2.9; trive 40; Mismatches 90; Indels

Query Match 6.0% Best Local Similarity 18.9% Matches 50; Conservative

6

Gaps

72;

Ouery Match 6.1%; Score 82; DB 21; Length 308; Best Local Similarity 19.9%; Pred. No. 1.6; Matches 41; Conservative 25; Mismatches 68; Indels 7

382 AA;

Sequence

```
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL16176-ABL30511), appressed DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL16176-ABL30511).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.Wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell
                                                                                 83 YQAYLQSRLGNYLPPMSQLL-TTARSWQACGHE-----PYQLPPEHLWGQIVPTLHLYQ 135
                                                                                                                                                                                136 DLKSRGI-LPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAID-----IWVPDL---EI 185
                           132 llssktplpttnylsafteslytdvqdsvskrletlqkdpattpfsvvfsdqlfhdplhp 191
                                                                                                                                               249 lesylesylslailqashilpkeatilrvtphdiepilppfsspesyl----iraihlyd 304
                                                                                                                                                                                                     --PIKTPSVLITKDKIGDHHTHE 52
                                                               ...----SHVGLQAHFETWLQMHHATKQEVVR 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Drosophila; developmental biology; cell signalling; insecticide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Disclosure; SEQ ID NO 39327; 21pp + Sequence Listing; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                               Drosophila melanogaster polypeptide SEQ ID NO 39327.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Myers EW;
                                                                                                                                                                                                                                                                                                                                                    ABB70845 standard; Protein; 1596 AA.
                                                                                                                                                                                                                                                               1: :: | | 359 cselsndvlsqllslfinqdsqqn 382
                                                                                                                                                                                                                                              186 KSQALYELQNRLCQYWLEHGENQN 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Li PWD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   23-MAR-2001; 2001WO-US09231.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23-MAR-2000; 2000US-191637P.
11-JUL-2000; 2000US-0614150.
                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Drosophila melanogaster.
17 LVACSAPIPTNPQVS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Adams M,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         WPI; 2001-656860/75.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (ABB57737-ABB72072)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (PEKE ) PE CORP NY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 1596 AA;
                                                          HDESV----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-PSDB; ABL14948
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200171042-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pharmaceutical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         interactions -
                                                                                                                                                                                                                                                                                                                                                                                                                  26-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27-SEP-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Venter JC,
                                                                                                                                                                                                                                                                                                                                                                                     ABB70845;
                          g
                                                                                     QQ
                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                ŏ
                                                                                                                                                                                                            g
                                                                                                                                                                                                                                             δ
```

```
1024 ea-----dhhqqqrelhqldqqqqqqaliladslphssssptssspptmpmpltti 1076
                                                                                                                                                                                                                                  1077 tapqll---plgpppphitstmpmpptmhmpimppppqcyqqlqp--ldptmsyhtiigs 1131
                                                                                      ---GNYLPPMSQLLTTA 105
                                                                                                                                                                                                   106 RSWQACGHEPYQLPPEHLWG--QIVPTLHL-----YQDLKSRGILPANTQIRSVYR 154
                                                                                                                                                                                                                                                                   155 NPELNQCAGGAAMSKHLTNSAIDIWVPDLEIKSQALYELQNRLCQYWLEHGENQNFG--- 211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human transferase; HTFS; agonist; antagonist; cellular signalling; proliferation; cell proliferative disorder; immune disorder; atherosclerosis; hepatitis; psoriasis; cancer; tumour; inflammation; AIDS; Addison's disease; allery; asthma; anaemia; cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus; multiple sclerosis; trheumatoid arthritis; pancratitis; systemic lupus erythematosus; thrombocytopenia; ulcerative colitis; haemodialysis; extracorporeal circulation; trauma; transgenic animal;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sequences AAB73494-AAB73535 represent novel human transferase proteins HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs encoding them. The proteins play important roles in the regulation of
                                            Gaps
                                                                 11 TLISSMLVACSAPIPTNPQVSPIKT--PSVLITKDKIGDHH-----THEHD 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               diagnosis, prevention and treatment of cell proliferative and immune system disorders and for identifying agonists and antagonists -
                                            68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bandman O, Patterson C;
 6.0%; Score 81; DB 22; Length 1596; 21.2%; Pred. No. 24; tive 38; Mismatches 109; Indels 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Human transferase polypeptides and polynucleotides useful for
                                                                                                                                  55 ESVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRL----
                                                                                                                                                                                                                                                                                                                                                                    1188 pvhtl----ttalhahqqg----gqqeaqtptl 1212
                                                                                                                                                                                                                                                                                                                                      212 ----LGLYATGAIHLDTQGFRKWGAQFSETNSI 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Lal P, Bar
Baughn MR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Human transferase HTFS-31, SEQ ID NO:31.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Claim 1; Page 126-127; 157pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB73524 standard; Protein; 448 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Yue H, Hillman JL,
Azimzai Y, Lu DAM,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           gene therapy; drug screening.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-NOV-2000; 2000WO-US30485.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           99US-0163595.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (INCY-) INCYTE GENOMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (first entry)
                                     Conservative
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2001-328796/34.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAH23831.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WO200132888-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           04-NOV-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       31-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10-MAY-2001.
                                 58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAB73524;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LL,
                                 Matches
                                                                                                 g
                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                    δλ
```

```
cellular signalling and proliferation. The HTFS proteins are useful for screening compounds for their effectiveness as agonists or antagonists of transferase activity, or for compounds that specifically bind to an HTFS protein or which modulates the activity of an HTFS protein.

Created compositions comprising an HTFS protein, HTFS paramaceutical compositions comprising an HTFS protein, HTFS protein are useful for treating a disease or condition associated with decreased or increased expression of functional HTFS pisorders with decreased or increased expression of functional HTFS pisorders of useful disporders. For example, diseases which may be treated using such compositions include cell proliferative disporders and immune disorders. For example, diseases which may be treated include atherosclerosis, hepatitis, psoriasis, cancers (including treated include atherosclerosis, hepatitis, psoriasis, cancers (including compasses, atopic dermatitis, diabetes mellitus, multiple sclerosis, chemmatoid arthritis, pancreatitis, sathma, anaemia, cirrhosis, cohors, confined the manual disease, and include and include and cancer, haemodialysis, extracorporeal circulation, confined to cancer, including lymphoma, leukaemia and manual manual diseases, included the natural or retains and manual proposed in manual diseases, for diagnostic purposes and cancers in manual the natural or retains and manual manual diseases, in compassing human diseases, and cancer in manual the natural or retains and manual proposed in manual diseases, and manual proposed and diseases and cancer, hence in manual the natural or retains and manual proposed in a manual diseases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            to generate hybridisation probes useful in mapping the naturally occurring genomic sequences. HTFS, and its catalytic or immunogenic fragments are useful for screening libraries of compounds in a variety of drug screening techniques. Antibodies which specifically bind HTFS may be used for the diagnosis of disorders associated with the expression of agonists, antagonists or inhibitors of HTFS. The present sequence represents an HTFS protein of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82 RYQAYLQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRG 141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      142 ILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDIWVPDLEIKSQALYELONRLCQYW 201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            65; Indels 30; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     32 ryreylqrq-----klfdtqhrvekmpdgsvalp---vlgetlpeqhl-qelrnr- 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human; chromosome mapping; gene mapping; gene therapy; forensic; food supplement; medical imaging; diagnostic; genetic disorder.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match

5.9%; Score 80; DB 22; Length 448;
Best Local Similarity 24.5%; Pred. No. 4.8;
Matches 40; Conservative 28; Mismatches 65; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    202 LEHG------ENQNFGLGLYATGAIHLDTQGFRKWG 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   135 grhgnllllsedcfqakqwknlgpelwetvalalgvqrlakrg 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Novel human diagnostic protein #12089.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ABG12098 standard; Protein; 233 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30-MAR-2001; 2001WO-US08631.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2000US-0540217.
2000US-0649167.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   WO200175067-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31-MAR-2000;
23-AUG-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            18-FEB-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11-0CT-2001.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ABG12098;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ABG12098
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ò
                     $$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$$
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
```

```
The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PGR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The and gene mapping are also used in diagnostics as expressed sequence tags of or identifying expressed genes. (I) is useful in gene therapy techniques (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a cod supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful in medical cod supplement. (II) and its binding partners are useful in medical cod supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating consists for general protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations controproposition and products dependent on DNA and amino acid sequences of data and products dependent on DNA and amino acid sequence data for this patent did not appear in the printed consisting the product of the this patent did not appear in the printed sequences of the this patent did not appear in the printed sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Neisseria meningitidis; Neisseria gonorrheae; antigen; vaccine; antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia; antibacterial; gene therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         118 LPPEHL-----WGQIVPTLHLYQDLK-SRGILPANTQIRSVYRNPELNQC---AGGAAMS 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        137 lspeeagehgrw--vpppsfgisdlegggadasrntsvygvwr--pilgcltglghtgtr 192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    77 mvlqtelrmdtpffrgpldrpqeesqllfpiqrpfqqqqccgllgdhfkpyssgsllhle 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5.8%; Score 78; DB 22; Length 233;
24.2%; Pred. No. 3;
tive 28; Mismatches 79; Indels 56; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  85 AYLQSRLGNYLP----PM-----SQLL-TTARSWQ---ACG-----HEPY-----Q 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20 mkletiilsklsgggktkhrmfsliefhlg---yhitwfslhhrkllldvtalgeviklg 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           33 IKTPSVLITKDKIGDHHTHEHDESVS-HVGLQAHFETWLQMHH------ATKQEVVRYQ 84
                                                                                                                                                                 New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Neisseria meningitidis ORF 280 protein sequence SEQ ID NO:1154.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   169 KHLTNSAIDIWVPDLEIKSQALYELQNRLCQYWLE 203
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  193 thltkhfvpwwrggl--rpreiwdlslsl---wle 222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                               Claim 20; SEQ ID No 42457; 103pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAY74840 standard; Protein; 304 AA
                                               Tang YT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    21-MAR-2000 (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    52; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity
                                                  Drmanac RT, Liu C,
                                                                                                    WPI; 2001-639362/73.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      233 AA;
(HYSE-) HYSEQ INC
                                                                                                                                N-PSDB; AAS76285.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY74840;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  д
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
AAG41671 standard; Protein; 388 AA.
                                                                                          (first entry)
                                                                                                                                                                                                                                termination sequence.
                                                                                                                                                                                                                                                                         Arabidopsis thaliana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           28-APR-1999;
30-APR-1999;
30-APR-1999;
06-MAY-1999;
06-MAY-1999;
07-MAY-1999;
11-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
18-MAY-1999;
20-MAY-1999;
21-MAY-1999;
21-MAY-1999;
22-MAY-1999;
22-MAY-1999;
23-MAY-1999;
23-MAY-1999;
25-MAY-1999;
25-MAY-1999;
25-MAY-1999;
25-MAY-1999;
25-MAY-1999;
25-MAY-1999;
25-MAY-1999;
25-MAY-1999;
                                                                                                                                                                                                                                                                                                                    EP1033405-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01-APR-1999;
06-APR-1999;
08-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                              25-FEB-2000;
                                                                                     18-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            19-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25-MAR-1999;
29-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  16-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       23-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              08-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10-JUN-1999;
                                             AAG41671;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA253015 to AA254536, AAA254577 to AA254615, and AAY74253 to AAX75941 represent novel Neisseria meningitis and N. gonorrheae polynucleotides and polypeptides. AAZ54571 to AAZ54576 and AAZ54616 to AAZ55473 represent polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent polypeptides, the polynucleotides, antibodies and composition. The polypeptides and compositions of immunogenic compositions. The polypeptides can be used in the manufacture of medicaments for treating or preventing infection due to Weisserial bacteria (e.g. meningitis and septicaemia), to detect the presence of Neisseria bacteria, or to raise antibodies. They may also be used to screen for agonists or antagonists, which may themselves have use as antibacterial agents. The polynucleotides of the invention may also be used in gene therapy protocols.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Scarselli M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     155 likadpegkvyyggrignygmglkkihsdagaafnavpaakrkvltghdafsymgkryhi 214
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --LPPEHLWGQIVPTLH----LYQDLKSRGILPANTQ----IRSVYR-NPELNQCAGGAA 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       74; Indels 52; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84 ------QAYLQSRLGNYLPPMSQL------LITARSWQACGHEPYQ------ 117
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       215 efiapggvsseaepsakgvaaiirgikregikavftenikdtrmvdriaketgvnvsgkl 274
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         30 VSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFE----TWLQ--MHHATKQEVVRY 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mora M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Novel Neisserial polypeptides predicted to be useful antigens for vaccines and diagnostics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Masignani V,
Scalato E, S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5.8%; Score 78; DB 21; Length 304; 21.1%; Pred. No. 4.5; Live 27; Mismatches 74; Indels :
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Fraser C, Galeotti C, Grandi G, Hickey E,
Petersen J, Pizza M, Rappuoli R, Ratti G,
Tettelin H, Venter JC;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Claim 2; Page 636-637; 1453pp; English.
                                                                                                                                                                                                                                                        98US-0099062.
98US-0103749.
98US-0103794.
                                                                                                                                                                                                                                                                                                                           98US-0103796.
                                                                                                                                            99WO-US09346
                                                                                                                                                                                                               98US-0094869
                                                                                                                                                                                          98US-0083758
                                                                                                                                                                                                                                                                                                                                                                                                                     (GENO-) INST GENOMIC RES.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       167 MSKHLTNSAIDIWV 180
    Neisseria meningitidis
                                                                                                                                                                                                                                                                                                                                                                                           (CHIR ) CHIRON CORP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WPI; 2000-062150/05.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  304 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-PSDB; AAZ53602.
                                             W09957280-A2
                                                                                                                                                                                                                              02-SEP-1998;
02-SEP-1998;
09-OCT-1998;
09-OCT-1998;
                                                                                                                                        30-APR-1999;
                                                                                                                                                                                                                                                                                                                           09-OCT-1998;
25-FEB-1999;
                                                                                             11-NOV-1999
```

275 ysdalgnapadtyi 288

118

RESULT 15 AAG41671

41;

g δλ g ŏ q δλ g

Query Match Best Local Si Matches 41;

Seguence

```
Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter;
                                   Arabidopsis thaliana protein fragment SEQ ID NO: 51877.
                                                                                                                                          9905 - 0125788
9905 - 0125788
9905 - 0126785
9905 - 0127462
9905 - 0128714
9905 - 0128714
9905 - 0130077
9905 - 0130691
9905 - 0130691
9905 - 0130691
9905 - 01310691
                                                                                                                                                                                                                                                        990S-0132863.
990S-0134256.
990S-0134218.
990S-0134219.
990S-0134370.
                                                                                                                                                                                                                                                                                                   990S-0134941.
990S-0135124.
990S-0135353.
990S-0135629.
990S-0136021.
                                                                                                             2000EP-0301439
                                                                                                                                                                                                                                 99US-0132484
                                                                                                                                                                                                                                             99US-0132486.
99US-0132487.
                                                                                                                                                                                                                                                                                                                                        99US-0136782
99US-0137222
                                                                                                                                                                                                                                                                                                                                                    99US-0137528
99US-0137502
                                                                                                                                                                                                                                                                                                                                                                      99US-0138094
                                                                                                                                                                                                                                                                                                                                                                            99US-0138540
                                                                                                                                                                                                                                                                                                                                                                                         99US-0139119
                                                                                                                                                                                                                                                                                                                                                                                                                                                99US-0139459.
                                                                                                                                                                                                                                                                                                                                                                99US-0137724
                                                                                                                                                                                                                                                                                                                                                                                               99US-0139452
                                                                                                                                                                                                                                                                                                                                                                                                    99US-0139453
                                                                                                                                                                                                                                                                                                                                                                                                            99US-0139492
                                                                                                                                                                                                                                                                                                                                                                                                                 99US-0139454
                                                                                                                                                                                                                                                                                                                                                                                                                        39US-0139455
                                                                                                                                                                                                                                                                                                                                                                                                                             99US-0139456
                                                                                                                                                                                                                                                                                                                                                                                                                                          99US-0139458
                                                                                                                                                                                                                                                                                                                                                                                   99US-0138847
                                                                                                                                                                                                                                                                                                                                                                                                                                    99US-0139457
```

US-0139461 US-0139462 US-0139763 US-0139763 US-0139981 US-0140354 US-0140354 US-0140892 US-0140892 US-0141893 US-0141893 US-0141893 US-0141893 US-0141893 US-0141893 US-0141893 US-0141893 US-0141893 US-0141893 US-0141893	990S-0142030. 990S-0142903. 990S-0142920. 990S-0143942. 990S-0143642. 990S-0144085. 990S-0144085. 990S-0144333. 990S-0144333. 990S-0144333. 990S-0144333. 990S-0144333.	903-014508 903-014508 903-014514; 903-014514; 903-0145213 903-0145213 903-014591; 903-014591; 903-014595 903-014595 903-014596 903-014595 903-014596 903-014596 903-014596	905 - 014730; 905 - 014741] 905 - 014741] 905 - 014743] 905 - 014817; 905 - 014834] 905 - 014856; 905 - 014956 905 - 014972 905 - 014972 905 - 014972
- JUN - 1999	06-JUL-1999; 08-JUL-1999; 08-JUL-1999; 13-JUL-1999; 14-JUL-1999; 14-JUL-1999; 16-JUL-1999; 16-JUL-1999; 19-JUL-1999; 19-JUL-1999; 19-JUL-1999; 20-JUL-1999; 20-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999; 21-JUL-1999;		5-AUG-199 5-AUG-199 9-AUG-199 9-AUG-199 1-AUG-199
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	7	: X X X X X X X X X X X X X X X X X X X	9 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8

```
77 KQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGH-EPYQLPPEHLWGQIVPTLHLYQ 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               136 DLKSR-GILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDIWVPDLEIKSQ----- 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       121 gqegrfsiyihpsrlrpvh-----isrhfs-----dreihsdhvtwgr 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -----GLGLY 215
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   87; Indels 114; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      31 SPIKTP-----SVLITKDKIGDHHTHEHDES---VSHVGLQAHFETWLQMH-HAT 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 5.8%; Score 78; DB 21; Length 388; Best Local Similarity 21.8%; Pred. No. 6.5; Matches 65; Conservative 32; Mismatches 87; Indels 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  990x - 0150566 .
990x - 015084 .
990x - 0151086 .
990x - 0151080 .
990x - 0151080 .
990x - 0151303 .
990x - 0153070 .
990x - 0153070 .
990x - 0153070 .
990x - 0154039 .
990x - 0154039 .
990x - 0154039 .
990x - 0155486 .
990x - 015548 .
990x - 015548 .
990x - 015929 .
990x - 016081 .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        99US-0161993.
    99US-0149930
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25-0CT-1999;
26-0CT-1999;
26-0CT-1999;
26-0CT-1999;
28-0CT-1999;
28-0CT-1999;
29-0CT-1999;
23-AuG-1999;
25-AuG 1999;
27-AuG-1999;
27-AuG-1999;
37-AuG-1999;
31-AuG-1999;
31-AuG-1999;
31-AuG-1999;
31-AuG-1999;
31-SEP-1999;
31-SEP-1999;
31-SEP-1999;
32-SEP-1999;
34-SEP-1999;
34-SEP-1999;
34-SEP-1999;
34-SEP-1999;
34-SEP-1999;
36-SEP-1999;
36-SE
                                                                                                                                                                                                                                                                                                                                                                                                                                 06-001-1999)
07-001-1999)
08-001-1999)
12-001-1999)
13-001-1999)
13-001-1999)
14-001-1999)
14-001-1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          21-OCT-1999;
21-OCT-1999;
21-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22-OCT-1999;
25-OCT-1999;
25-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14-OCT-1999;
14-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 21-0CT-1999;
21-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           22-0CT-1999;
22-0CT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18-0CT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        21-0CT-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
        οp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      οy
```

- 159 ismvdaerrllanaledpdnghfvllsesciplhtfdytyrylmhanvsfidsfedlgph 218 οqα
  - Db δy

Search completed: July 30, 2002, 16:15:08 Job time: 481 sec

		, 1
		r.
		•

```
July 30, 2002, 16:07:52; Search time 13.04 Seconds (without alignments) 468.282 Million cell updates/sec
                                                                                                                                                                                                           1355
1 MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250
                                                                                                                                                                                                                                                                                                                                                                      231628
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                        Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                          231628 seqs, 24425594 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                      OM protein – protein search, using sw model
                                                                                                                                                                                                                                                                   BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                           US-09-674-779-2
                                                                                                                                                                                     Title:
Perfect score:
                                                                                                                                                                                                                                                                 Scoring table:
                                                                                                                                                                                                                             Sequence:
                                                                                                                                                                                                                                                                                                                          Searched:
                                                                                                              Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. 

Database :

SUMMARIES

Description	Sequence 2, Appli	ì	Sequence 2, Appli	Sequence 2, Appli	4 0	٠ ١	Sequence 2, Appli	, ,	Sequence 1, Appli	2 2	,	Sequence Z, Apput	, u	, v	, (	Sequence 2, Appli	Datont No 5177107			4 0	4 (	ų c	7	Sequence 22, Appl	$\sim$	15. 7	-	3, 18
ID	US-08-458-120-2	US-08-867-970-2	-09-326	-08-972-902-	-09-071-101-	US-09-369-618-2			US-08-588-258B-40	896-08295-4	٠,	-439-313-	08-986-485-5	-844-085-	09-353-332-	09-149-92	,	5177197-30	US-08-660-347-2	US-08-700-6704-23	-08-596-366-2	US-08-967-104-2	7 107 107 00	77-118-6/8-80-	-08-828-2	US-09-286-691-15	∹	US-08-893-852A-3
DB	ч	7	4	m	m		m	4		Ŋ	7	4	c	~	7	4	9	9	m	N	N	0	• (	n .	4	4	4	m
Query Match Length	578	578	578	368	755	755	755	462	420	420	1141	140	1091	305	437	512	410	1394	436	473	883	883		777	526	473	473	657
Query		5.8	5.8	5.7			5.5		5.4	5.4	5.4		5.2	5.2	5.2	5.2	5.1	5.1	5.1	5.1	5.1	5.1	ı с	٠	o. c		5.0	5.0
Score	78	78	78	77	75	75	75	74	73	73	73	71.5	71	70.5	70	7.0	69	69	68.5		68.5	68.5	8	3 6	99	68	68	99
Result No.	1	7	m	4	S	9	7	œ	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	1	77	25	56	27

nce 3, Appli nce 2, Appli nce 2, Appli nce 2, Appli nce 17, Appli nce 17, Appli nce 17, Appli nce 3, Appli nce 6, Appli nce 6, Appli nce 1, Appli nce 2, Appli nce 2, Appli nce 2, Appli nce 2, Appli nce 2, Appli nce 2, Appli	
sednence Sed	9. 0
US-08-821-818-3 US-08-464-227C-2 US-08-455-146-2 US-08-468-5065-2 US-08-468-5768-17 US-08-468-5778-17 US-09-173-914-2 US-09-173-914-2 US-09-102-248-3 US-09-102-248-3 US-09-102-248-3 US-09-102-248-3 US-09-102-248-3 US-09-102-248-3 US-08-411-389-2 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1 US-08-510-284-1	S/08458120 LL.  The control of the c
5.0 4472 2 5.0 4472 2 5.0 622 2 5.0 622 2 5.0 712 2 5.0 712 2 5.0 712 2 712 2 712 2 712 2 713 2 714 2 714 2 715 2 716 2 717 2 718 2 719 2	ion Ug  i, ET A  in Hu  ES: 4  DRESS: 4  DRESS: 5  ELLA, 5  ELLA, 6  ELA, 6  ELLA, 6
68 68 68 67 67 67 67 67 68 67 67 67 67 67 67 67 67 67 67	1 10-2 ence 2, A ence 2, A ERAL INFO PPLICANT: ITLE OP PPLICANT: ITLE OF UMBER OF ORRESPOND ORRESPOND ORRESPOND ORRESPOND ORREST: CITY: R STRET: COUNTRY: CO
99.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.0.	RESULT US-08-458 Sequenc Releach APPL APPL APPL CORR CORP CORP COMPI COM

11; 71; Indels 58; Gaps Ouery Match 5.8%; Score 78; DB 1; Length 578; Best Local Similarity 24.3%; Pred. No. 3.6; Matches 49; Conservative 24; Mismatches 71; Indels

STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-458-120-2

82 RYQ----AYLQ-----SRLGNYLPPMSQLL-----TTARSWQACGHEPYQLPPEHLWGQIV 128

δλ

```
58; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    129 PTLH-----LYQDLKSRGILPANTQIRSVYRNPEL-NQC---AGGAAMSKHLTNSAI 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82 RYQ----AYLQ-----SRLGNYLPPMSQLL-----TTARSWQACGHEPYQLPPEHLWGQIV 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         352 -TLREKMKSNEAASKELASKKSLPMNYYTVFYHVQEQLPRDCFVVSEGA-----NTM 402
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           294 RYOPDVKFIQVDICAEELGNNVKPAVTLLGNIHAVTKQLLEELDKTPWQYPPESKWWK-- 351
1 : | | : | : | : : | 352 -TLREKMKSNEAASKELASKKSLPMNYYTVFYHVQEQLPRDCFVVSEGA------NTM 402
                                                                                        177 DIWVPDLEIKSQALYELQNRLCQYWLEHGENQNFGLGL---YATGAIHLD-----T 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5.8%; Score 78; DB 4; Length 578; 24.3%; Pred. No. 3.6; tive 24; Mismatches 71; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                  Human Oxalyl-CoA Decarboxylase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: Human Genome Sciences, Inc. STREET: 9410 Key West Avenue CITY: Rockville
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 02-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US94/05561
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/867,970
FILLING DATE: JUN-03-97
APPLICATION NUMBER: US 08/458,120
FILLING DATE: 02-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/326,217
                                                                                                                                                                                                                                                                      US-09-326-217-2
Sequence 2, Application US/09326217
Patent No. 6200796
GENERAL INFORMATION:
APPLICANT: Olsen, Henrik S.
APPLICANT: Application Timothy
APPLICANT: Adams, Mark D.
                                                                                                                                                 225 QGFRKWGAQFSETNSICRHVLP 246
                                                                                                                                                                        MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   30,446
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELECOMMUNICATION INFORMATION:
TELEPHONE: 301-309-8504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO: 2: SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 18-MAY-1994 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: Benson, Robert H. REGISTRATION NUMBER: 30, REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              : 578 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 5.8%
Best Local Similarity 24.3%
Matches 49; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301-309-8439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    , MOLECULE TYPE: protein US-09-326-217-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                    TITLE OF INVENTION: HUNDMER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   OPERATING SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ns
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Qγ
                                                                                                                                                   δà
                                                                   QΥ
                                                                                                          αq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  129 PŢLH-----LYQDLKSRGILPANŢQIRSVYRNPEL-NQC---AGGAAMSKHLTNSAI 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    82 RYQ---AYLQ----SRLGNYLPPMSQLL----TTARSWQACGHEPYQLPPEHLWGQIV 128
                                                                                               129 PTLH-----LYQDLKSRGILPANTQIRSVYRNPEL-NQC---AGGAAMSKHLTNSAI 176
                                                                                                                                                   177 DIWVPDLEIKSQALYELQNRLCQYWLEHGENQNFGLGL---YATGAIHLD-----T 224
                                                                                                                                                                                 Query Match 5.8%; Score 78; DB 2; Length 578; Best Local Similarity 24.3%; Pred. No. 3.6; Matches 49; Conservative 24; Mismatches 71; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: Olsen, Henrik S.
APPLICANT: Oleman, Timothy
APPLICANT: Adams, Mark D.
TITLE OF INVENTION: Human Oxalyl-CoA Decarboxylase
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/867,970
FILLING DATE: UNN-3-97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEE: Human Genome Sciences, Inc.: 9410 Key West Avenue
Rockville
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICATION NUMBER: WO PCT/US94/05561
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/458,120
FILING DATE: 02-JUN-1995
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER: IBM PC compatible OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/08867970 Patent No. 5945273
                                                                                                                                                                                                                                       225 QGFRKWGAQFSETNSICRHVLP 246
                                                                                                                                                                                                                                                                 TELECOMMUNICATION INFORMATION: TELEPHONE: 301-309-8504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME: Benson, Robert H. REGISTRATION NUMBER: 30, REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           : 578 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 301-309-8439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; MOLECULE TYPE: protein US-08-867-970-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20850
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ξ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                 US-08-867-970-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Qγ
                                                                                                                g
                                                                                                                                                         δχ
                                                                                                                                                                                                  g
                                                                                                                                                                                                                                            δλ
                                                                                                                                                                                                                                                                                   g
```

---REWVATLPIGYADGLIRH 289

263 YGATYTATE-----

g

```
6
                                     403 DI------GRTVLQNYLPRHRLDAGTFGTMGVGLGFAIAAAVVAKDRSPGQWIICV 452
177 DIWVPDLEIKSQALYELQNRLCQYWLEHGENQNFGLGL---YATGAIHLD-----T 224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         50 THEHDESV---SHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTT-- 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               90 TRKEDANLAAKNHISLTVFREDWLE--NLTLEATLRIHLKVDSGWGRL-----GIRTTEE 142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                143 ARRIEATSTNDHQLQLEGIYTHFATADQLETSYFEQQLAKFQTILTSLKKRPTYVHTANS 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             165 AA--MSKHLTNSAIDIWV-----PDLEIKSQALYELQNRLCQYW-----LEHGENQN 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         105 ARSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        98; Indels 38; Gaps
                                                                                                                                                                                                                                                                                     APPLICANT: Frankel, Fred R.
APPLICANT: Portnoy, Daniel A.
TITLE OF INVENTION: BACTERRAL VACCINE VECTOR AND METHODS OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
5.7%; Score 77; DB 3; Length 368;
Best Local Similarity 21.3%; Pred. No. 2.3;
Matches 46; Conservative 34; Mismatches 98; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                     ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C. STREET: One Commerce Square, 2005 Market Street, 22nd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OPERATING SYSTEM: PC_DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              210 FGLGLYATGAIHLDTQGFRKWGA--QFSETNSICRH 243
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICATION NUMBER: US/08/972,902
FILING DATE: 18-NOV-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9596-51
                                                                                                                                                                                                                          Sequence 2, Application US/08972902
Patent No. 6099848
GENERAL INFORMATION:
                                                                                 225 QGFRKWGAQFSETNSICRHVLP 246
                                                                                                                         453 EGDSAFGFSGMEVETICRYNLP 474
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER: IBM PC COMPATIBLE OPERATING SYSTEM: PC-DOS/MS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATTORNEY ACTOR NAME: DOYLE-LEATY, KATHIYN REGISTRATION NUMBER: 36,317
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 95.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-965-1284
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TELEX: 831-494
INFORMATION FOR SEQ ID NO: 2:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       368 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELEPHONE: 215-500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                           NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY: U.S.A.
ZIP: 19103-7086
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; MOLECULE TYPE: protein US-08-972-902-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ': Fl.
Philadelphia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                       RESULT 4
US-08-972-902-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STREET:
                                                                                   δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             òγ
```

```
340 STLTTTYSSGVATDPQI-----VTTDL---HHGCTDQHTGTSASAPLAAGMIALAL 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----RLGNYLPPMSQLLTTA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                388 EANPFLTWRDMQHL----VVRASKPAHLQAEDWRTNGVGRQVSHHYGYGLLDAGLLVDTA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            106 RSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSR--GILPANTQIRSVYRNPELNQCAG 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       444 RTW------IVIRENVSACAG 481
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     164 GAAMSKHLTNSAIDIWVPDLEIKSQALYELQNRLCQYWLEHGE----NQNFGLG----LYA 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ----RSLEHVQAQLTLSYSRRGDLEISLTSPMGTRSTLVA 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  14 SSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHH--THEHDESVSHVGLQA----- 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ouery Match 5.5%; Score 75; DB 3; Length 755; Best Local Similarity 22.0%; Pred. No. 12; Matches 56; Conservative 32; Mismatches 72; Indels
                                                                                                       APPLICANT: Lok, Si
APPLICANT: Lok, Si
APPLICANT: Jaspers, Stephen R.
TITLE OF INVENTION: HUMAN PROHORMONE CONVERTASE
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                               OPERATING SYSTEM: DOS
SOFTWARE: Fastsep for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,101
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65 ---HFETWLQMHHATKQEVVRYQ--AYLQS----
                                                                                                                                                                                                                ADDRESSEE: ZymoGenetics, Inc. STREET: 1201 Eastlake Avenue East
                                           Sequence 2, Application US/09071101 Patent No. 6013503
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: 97.
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-442-6673
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31,648
                                                                                                                                                                                                                                                                                                                                                                                             IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE CHARACTERISTICS:
LENGTH: 755 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                          Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: Parker, Gary E. REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEFAX: 206-442-6678
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOLECULE TYPE: protein FRAGMENT TYPE: internal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CLASSIFICATION:
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              482 -----LHNSI-----
                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        217 TGAIHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 523 IRPLDVSTEGYNNW 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                    GENERAL INFORMATION:
                                                                                                                                                                                                                     ADDA.
STREET: 120.
TTTY: Seattle
                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                                                                                                                                                                                                                      MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE:
                                                                                                                                                                                                                                                                                                   COUNTRY: U:
ZIP: 98102
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE
                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY:
  RESULT 5
US-09-071-101-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-071-101-2
RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Dβ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
```

us-09-674-779-2.rai

```
164 GAAMSKHLTNSAIDIWVPDLEIKSQALYELQNRLCQYWLEHGE----NQNFGLG----LYA 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -----LHNSI------RSLEHVQAQLTLSYSRRGDLEISLTSPMGTRSTLVA 522
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           65 ---HFETWLQMHHATKQEVVRYQ--AYLQS-------RLGNYLPPMSQLLTTA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        388 EANPFLTWRDMOHL----VVRASKPAHLQAEDWRTNGVGRQVSHHYGYGLLDAGLLVDTA 443
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             106 RSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSR--GILPANTQIRSVYRNPELNQCAG 163
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         340 STLTTTYSSGVATDPQI----VTTDL---HHGCTDQHTGTSASAPLAAGMIALAL 387
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          72; Indels 94; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14 SSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHH--THEHDESVSHVGLQA----- 64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 5.5%; Score 75; DB 3; Length 755; Best Local Similarity 22.0%; Pred. No. 12; Matches 56; Conservative 32; Mismatches 72; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GENERAL INFORMATION:
APPLICANT: Lok, Si
APPLICANT: Jaspers, Stephen R.
TITLE OF INVENTION: HUMAN PROHORMONE CONVERTASE 4
FILE REFERENCE: 97-05D1
                                                                                                                      APPLICANT: Jaspers, Stephen R.
TITLE OF INVENTION: HUMAN PROHORMONE CONVERTASE 4
FILE REFERENCE: 97-05D2
                                                                                                                                                                                      CURRENT APPLICATION NUMBER: US/09/369,618
CURRENT FILING DATE: 1999-08-06
EARLIER APPLICATION NUMBER: US 09/071,101
EARLIER FILING DATE: 1998-05-01
EARLIER FILING DATE: 1998-05-06
EARLIER FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION NUMBER: US/09/369,617
CURRENT FILING DATE: 1999-08-06
EARLIER PEDLICATION NUMBER: US 09/071,101
EARLIER FILING DATE: 1998-05-01
EARLIER APPLICATION NUMBER: US 60/044,015
BARLIER FILING DATE: 1998-05-06
NUMBER OF SEQ ID NOS: 16
SOFTWARRE: PRACESO FOR Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 2, Application US/09369617 Patent No. 6127162
               Sequence 2, Application US/09369618 Patent No. 6100041
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      217 TGAIHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   523 IRPLDVSTEGYNNW 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-369-617-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Homo sapiens
                                                                       GENERAL INFORMATION:
APPLICANT: Lok, Si
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ ID NO 2
LENGTH: 755
                                                                                                                                                                                                                                                                                                                                                                                                 SEQ ID NO 2
LENGTH: 755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       US-09-369-617-2
US-09-369-618-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-09-369-618-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                   TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
```

```
15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          117 VSENHMRKKLYENGVTDSLKSNFALLLKLSEELLDKWLSYPETQHVPLSQHMLGFAMKSV 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   113 HEPYQLPPEHLW--GQIVPTLHLYQDLK-----SRGILPANTQIRSVYRNPELNQCAGGA 165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     62 HERYG-PVVSFWFGRRLVVSLGTVDVLKQHINPNKTLDPFETMLKSLLR----YQSGGGS 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                166 AMSKHLTNSAIDIWVPD-LEIKSQALYELQNRLCQYWLEHGENQNFGLGLYATG-AIHLD 223
                                                                                                                                                                                                                               388 EANPELTWRDMOHL----VVRASKPAHLQAEDWRTNGVGRQVSHHYGYGLLDAGLLVDTA 443
                                                                                                                                                                                                                                                                                                                      444 RIW-----INTQ-----PQRKCAVRQSRPTPILPL-----IXIRENVSACAG 481
                                                                                                                                                                                                                                                                                                                                                                            164 GAAMSKHLTNSAIDIWVPDLEIKSQALYELQNRLCQYWLEHGE---NQNFGLG----LYA 216
                                                                                                                                                                                                                                                                                                                                                                                                                   ----RSLEHVQAQLTLSYSRRGDLEISLTSPMGTRSTLVA 522
                                                                                                                              65 ---HFETWLQMHHATKQEVVRYQ--AYLQS-------RLGNYLPPMSQLLTTA 105
                                                                                                                                                                                                                                                                                  106 RSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSR--GILPANTQIRSVYRNPELNQCAG 163
                                                   Gaps
                                                                                              14 SSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHH--THEHDESVSHVGLQA----- 64
                                                94;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match 5.5%; Score 74; DB 4; Length 462; Best Local Similarity 25.2%; Pred. No. 7.2; Matches 35; Conservative 21; Mismatches 67; Indels
y Match 5.5%; Score 75; DB 3; Length 755; Local Similarity 22.0%; Pred. No. 12; hes 56; Conservative 32; Mismatches 72; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Henry Yue
APPLICANT: Karl J. Guegler
APPLICANT: Neil C. Corley
APPLICANT: Janice Au-Young
TITLE OF INVENTION: HUMAN CYTOCHROME P450
FILE REPERENCE: PP-0608 US
CURRENT APPLICATION NUMBER: US/09/165,241
CURRENT FILING DATE: 1998-10-01
NUMBER OF SEQ ID NOS: 2
SEQ ID NO I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 40, Application US/08588258B Patent No. 5929207 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Sequence 1, Application US/09165241; Patent No. 6130077
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            224 TOGFRKWGAQFSETNSICR 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   177 TQ--MVMGSTFEDDQEVIR 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; OTHER INFORMATION: 991729
US-09-165-241-1
                                                                                                                                                                                                                                                                                                                                                                                                                               482 ------THNSI------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           217 TGAIHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      523 IRPLDVSTEGYNNW 536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  US-08-588-258B-40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 8
US-09-165-241-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LENGTH: 462
            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURE:
                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                               g
                                                                                                                                                                                               δy
                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                         δλ
                                                                                                                                                                                                                                                                                                                                   qq
                                                                                                                                                                                                                                                                                                                                                                                   δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
```

```
153 PVP-----GCVNTTEVDIKK---SSRMRNPHKTRKSVYGLONDIRSHSPTHTPTPETKPP 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     76 TKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSW--QACGHEPYQLPPE--HLWGQIVPTL 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               132 HLYQDLKSRGILPANTQIR--SVYRNPELNQCAGGAAMSKHL----TNSAIDIWVPDLEI 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        23 PIPTINPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMH-----HA 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5.4%; Score 73; DB 2; Length 420; 20.5%; Pred. No. 8.1; tive 40; Mismatches 102; Indels
APPLICANT: H. RODELT HOLVILZ Et al.
TITLE OF INVENTION: REGULATORS OF G-PROFEIN SIGNALLING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT: Massachusetts Institute of Technology
TITLE OF INVENTION: REGULATORS OF G-PROTEIN SIGNALLING
NUMBER OF SEQUENCES: 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          186 KSQALYELQNRLCQYWLEHGENQNFGLGLYATGAIHLDTQGFRK 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                39,109
FR: 01997/216001
                                                                                                                                                                                  COMUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 40, Application PC/TUS9608295 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                         ATTORNEY/AGENT INFORMATION:
NAME: Bleker-Brady, Kristina
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 01997,
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617.428-0200
TELEFAX: 617.428-7045
                                                                               E: Clark & Elbing LLP
176 Federal Street
                                                                                                                                                                                                                                                                                                                        January 12, 1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ ID NO: 40:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           46; Conservative
                                         NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-588-2588-40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ZIP: 02110-2804
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local Similarity
                                                                                                                                                                                                                                                                                                                                            CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      02110-2804
                                                                                                                                                       USA
                                                                                                                     Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                          FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Boston
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        USA
                                                                                                                                                                       02110
                                                                               ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PCT-US96-08295-40
                                                                                           STREET:
                                                                                                                                                     COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ద
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
```

```
153 PVP-----GCVNTTEVDIKK---SSRMRNPHKTRKSVYGLQNDIRSHSPTHTPTPETKPP 204
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             76 TKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSW--QACGHEPYQLPPE--HLWGQIVPTL 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         132 HLYQDLKSRGILPANTQIR--SVYRNPELNQCAGGAAMSKHL----TNSAIDIWVPDLEI 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   265 WELEASKE----PSQORVKRWGFGMDEALKDPVGREOFLKFLESEFSSENLRFWLAVEDL 320
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ; Pred. No. 8.1;
40; Mismatches 102; Indels 36; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23 PIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMH------HA 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5.4%; Score 73; DB 5; Length 420; 20.5%; Pred. No. 8.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             186 KSQALYELQNRLCQYWLEHGENQNFGLGLYATGAIHLDTQGFRK 229
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MEDIUM TYPE: Floppy dlsh
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08295
"TT NG DATE: 31-MAY-1996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Zon, Leonard and Richardson, Paul
TITLE OF INVENTION: The Gene and Uses Thereof
NUMBER OF SEQUENCES: 2
                                                                                                                                                                                                                                          CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Bleker-Brady, Kristina
REGISTRATION NUMBER: 39,109
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-8070
TELEFAX: 617/542-8070
                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/588,258
FILING DATE: 12-JAN-96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 2, Application US/08363300 Patent No. 5700927 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
                                                                                                                                                                                                                                                                                                                                                                                           TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 40:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        not relevant
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               420 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE CHARACTERISTICS:
LENGTH: 420 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         46; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: linear
MOLECULE TYPE: protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best Local Similarity
Matches 46; Conserva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TYPE: amino acid STRANDEDNESS: no
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY: USA
ZIP: 02110-2804
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDALL
STREET: 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PCT-US96-08295-40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-08-363-300-2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     qq
```

US/08/363,300

```
467 TA----TCAH-----PESLKGQSIFSVLPDSFVCDDFPKPQIITQPETTMAVVGKDIRF 516
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   160 QCAGGAAMSKHLTNSAIDIWVPDLEIKSQALYELQNRLCQYWLEHGENQNFGLGLYATGA 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               517 TCSAASSSSPMTEA----WKKDNEVLANA--DMEN-FAHVRAQDGEVMEY-----TTI 563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    104 TARSWQACGHEPYQLPPEHLWGQ----IVPTLHLYQDLKSRGILPANTQIRSVYRNPELN 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           55 ESVSHVGL-----QAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPP-----MSQLLT 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       408 ESLEHLNLGENAIRSVQFDAFAKMKN-LKELYISSESFLCDCQLKWLPPWLMGRMLQAFV 466
106 RSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGA 165
                                               ----LNYPAGGT 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match 5.2%; Score 71; DB 3; Length 1091; Best Local Similarity 20.6%; Pred. No. 62; Matches 45; Conservative 35; Mismatches 88; Indels 5
                                                                                                                                                                                                                                                                                                                 GENERAL INFORMATION:
APPLICANT: WU, SHUJIAN
APPLICANT: SWEET, RAYMOND
APPLICANT: TRUNEH, ALEMSEGED
TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              220 IHLDTQGF------RKWGAQFSETNSICRHVLP 246
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     UMBER: US/08/986,485
08-DEC-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GH-70264
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/059,448
FILING DATE: 22-SEP-1997
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                             ; Sequence 5, Application US/08986485; Patent No. 6046030
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           REFERENCE/DOCKET NUMBER: GH
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TELEX: 846169
INFORMATION FOR SEQ ID NO: 5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE CHARACTERISTICS:
LENGTH: 1091 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  E: Diskette
IBM Compatible
                                                    106 R----HHHADTPP--LWCR-
                                                                                                     166 AMSKHLTNSAIDIWV 180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; TOPOLOGY: linear; MOLECULE TYPE: protein US-08-986-485-5
                                                                                                                                                127 AVA----YSCLSDWL 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER:
FILING DATE: 08-DEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER: IBM COM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  19482
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ΡA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COUNTRY:
                                                                                                                                                                                                                                                US-08-986-485-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
                                                         g
                                                                                                                                                     qq
                                                                                                        δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             50 HHTHEHTDTLPYGHWHTHCHTVTWTHLHTITPPHTLP----VDTRTHRHCHTDTQNTVTR 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        48 HHTHEHDESV--SHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTA 105
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         64 AHFETWLQ---MHHATKQEVV----RYQAYLQSRLGNYLPPMSQLLTTAR------ 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    13 ISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEH---DESVSHVG-----LQ 63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         43; Indels 49;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Solk, John
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: DAGNOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REPERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 477
LENGTH: 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ouery Match
Best Local Similarity 21.5%; Pred. No. 2.1;
Matches 29; Conservative 14; Mismatches 43; Indels 4
                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match

5.4%; Score 73; DB 1; Length 1141;
Best Local Similarity 25.2%; Pred. No. 39;
Matches 34; Conservative 16; Mismatches 49; Indels
                                                                                                                                   04590/009001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GENERAL INFORMATION:
APPLICANT: Xu. Jiangchun
APPLICANT: Xu. Jiangchun
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Harlocker, Susan Louise
APPLICANT: Hand year
APPLICANT: Reed, Steven G.
APPLICANT: Ranger, Gary
APPLICANT: Retter, Mark
APPLICANT: Retter, Mark
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 477, Application US/09439313
Patent No. 6329505
                                                                                                                            REFERENCE/DOCKET NUMBER: 0459.
TELECOMMUNICATION INFORMATION:
TELEFAN: (617) 542-8070
TELEFAX: (617) 542-806
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                              ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
                                                                                                                                                                                                                                                                                   LENGTH: 1141 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 107 ----- SWQACGHEP 115
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              671 RGRRHASFESCGKRP 685
                                                                                                                                                                                                                                                                                                                                                 , MOLECULE TYPE: protein US-08-363-300-2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TYPE: PRT
ORGANISM: Homo sapiens
                                                                                                                                                                                                                                                                                                     TYPE: amino acid
TOPOLOGY: linear
                                          CLASSIFICATION:
                      FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-439-313-477
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-439-313-477
```

В ò

δy

δy

```
Search completed: July 30, 2002, 16:15:28
Job time: 456 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         130 PLFDPKTSSTYKDVSCSS-----
                                                                                                                                                                                                                                              ; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-353-332-2
          APPLICANT: Lamb, Christopher
                                                                                                                                                                                                                               437
                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Op
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              133 -----LY-QDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDIWVPDL 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 87 PSPSNIQELYLESLEKLGINPLEHDIRFVEDNWE-NPSTGSAGL-----GWEVWLDGM 138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           184 EIKSQALYELQNRLCQYWLEHGENQNFGLGLYATGAIHLD-TQGFRKWGAQFSETNSI 240
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           :| ||| : : | | : : | :|:
----VGGLATGPVTAEVTYGLERLASYIQEVDSV 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   47;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match 5.2%; Score 70.5; DB 1; Length 305; Best Local Similarity 22.5%; Pred. No. 9.4; Matches 40; Conservative 23; Mismatches 68; Indels 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  84 QAYLQSRLGNYLPPMSQLLTTA-RSWQACGHEPYQLPPEHLWGQIVPTLH--
:|| : || : :|| 564 LHLRHUTFGHEGRYQCIITNHFGSTYSHKARLTVNVLP 601
                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/844,085
FILING DATE: 18-APR-1997
CLASSIFICATION: 435
                                                                                                                                                                        APPLICANT: Lawlor, Elizabeth
TITLE OF INVENTION: No. 5756330el Compounds
NUMBER OF SEQUENCES: 12
                                                                                                                                                                                                                             CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation STREET: 709 Swedeland Road
CITY: King of Prussia
                                                                                                                   Sequence 6, Application US/08844085
Patent No. 5756330
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRIOR APPLICATION DATA:
APPLICATION NUMBER: 9608002.3
FILING DATE: 18-APR-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 15
US-09-353-332-2
; Sequence 2, Application US/09353332
; Patent No. 6316697
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38,891
                                                                                                                                                                                                                                                                                                                                                    ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM COMPATIBLE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
LENGTH: 305 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GENERAL INFORMATION:
APPLICANT: Dixon, Richard A.
APPLICANT: Xia, Yiji
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME: Gimmi, Edward R
REGISTRATION NUMBER: 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRANDEDNESS: single
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; MOLECULE TYPE: protein US-08-844-085-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   || :| | |
139 EI-TQFTYFQQ----
                                                                                                                                                                                                                                                                                                                                COUNTRY: USA
ZIP: 19406-0939
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          linear
                                                                                                                                                       GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                  ΡA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY:
                                                                              RESULT 14
US-08-844-085-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δλ
```

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   85 AYLQSRLGNYL-----PPM-----SQLLTTARSWQACGHEPYQLPPEHLWGQIV 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          129 P----TLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDIWVPDL 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25 PTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATKQEVVRYQ 84
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        96;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match 5.2%; Score 70; DB 4; Length 437; Best Local Similarity 19.5%; Pred. No. 19; Matches 43; Conservative 21; Mismatches 61; Indels
TITLE OF INVENTION: CONSTITUTIVE DISEASE RESISTANCE (CDR1)
TITLE OF INVENTION: GENE AND METHODS OF USE THEREOF
FILE REPERENCE: SALK2820-1
CURRENT APPLICATION NUMBER: US/09/353,332
CURRENT FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,696
BARLIER APPLICATION NUMBER: 60/092,696
BARLIER FILING DATE: 1998-07-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSEQ for Windows Version 4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -----SQCTA----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          153 -LENQASCSTNDNTCSYSLSYGDNS----YTKGNIAVDT 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   184 EIKSQALYELQNRLCQYWLEHGENQNFGLGLYATGAIHLDT 224
```

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

July 30, 2002, 16:08:27 ; Search time 19.09 Seconds (without alignments) 1258.373 Million cell updates/sec Run on:

1 MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250 US-09-674-779-2 Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5

Scoring table:

Searched:

283138 Total number of hits satisfying chosen parameters:

283138 seqs, 96089334 residues

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

PIR\_71:\* |: pir1:\* pir2:\* pir3:\* pir4:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description	todacount boundario		hypothetical prote				vefs protein predu	hypothetical profe	hypothetical profe	probable exodensor	phosphoribosylform	probable oxidored	hypothetical prote				hypothetical prote	probable membrane	probable linoprote	glaculte inpopiote	gircie membrane-bound lut	membrane-bound lyt		exchinctes of the state of the	probable periplem	hypothetical proto		hynothetical proto	Ω
SUMMARIES	ID	G87436	AG3019	B98265	AH3510	AH0644	AD2215	F64855	C90815	G85674	D71490	AI1295	A70523	G83424	E36792	B82162	847290	164172	S70295	AI0245	B83821	A98088	A65064	C85933	AD0865	E81923	G86325	T28662	G82697	JC6042
	DB	2	7	~	7	7	~	7	7	7	7	7	~	7	7	~	~	7	7	~	7	~	7	~	7	7	7	7	7	7
	Query Match Length	216	587	634	637	263	889	320	320	320	1004	739	536	532	599	373	511	209	528	207	297	365	365	365	1123	308	1062	464	467	474
ф	Query Match	7.5	7.5	•	•	•	•	•	•	•				•								6.1						•	5.9	
	Score	101	101	101	96.5	93	90	88	88	88		87.5	98	84.5		(*)	83.5	83	83	(1	(1	82.5	$\alpha$	$^{\alpha}$	$^{\circ}$	82	81.5	80.5	80.5	80.5
	Result No.	Н	7	m	₹	'n	9	7	σ,	ָס	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	29

phosphoribosyl form	TRNA isonenteny' +	membrane-bound 1 111	adhosin probablo	hynothetical proto	Conserved hypothet	ferredoxinnitrit	acyl-CoA thioester	transcription ream	L1 protein - bovin	reversed polarity	hypothetical profe	TEXT TOWN TOWN TOWN TOWN	Dantothenate kinas	aldose 1-asimes	cysteinyl-tRNA syn
AG1667	T48100	AF0864	D81182	B83372	C82217	S16603	B83153	T34829	PlWLB4	A54282	S67787	AE0478	C84009	T07719	B84266
7	a	^1													
			C	( 1	C	7	7	2	7	7	7	~	~	~	7
739	330	365	304 2	458 2	492 2	594 1	289 2	350 2	506 1	612 2	649 2	878 2	316 2	358 2	494 2
5.9 739	5.9 330	5.8 365	5.8 304 2	5.8 458 2	5.8 492 2	5.7 594 1	5.7 289 2	5.7 350 2	5.7 506 1	5.7 612 2	5.7 649 2	5.7 878 2	5.6 316 2	5.6 358 2	5.6 494 2
5.9	79.5 5.9 330	5.8	78 5.8 304 2	78 5.8 458 2	78 5.8 492 2	5.7	5.7	77 5.7 350 2	5.7	77 5.7 612 2	77 5.7 649 2	77 5.7 878 2	5.6	76.5 5.6 358 2	5.6

## ALIGNMENTS

RESULT 1 G87436
conserved hypothetical protein CC1512 [imported] - Caulobacter creacentus
C; Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C;Accession: G87436
R, Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.I.; Haff, D.H. KO
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.I.; Shaniro, I
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141 2001
A; Title: Complete Genome Sequence of Caulobacter crescentus
A; Reference number: A87249; MUID:21173698; PMID:11259647
A; Accession: G87436
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-216 <sto></sto>
A)Cross-references: GB:AE005673; NID:q13422891; PIDN.AAK23491 1. Copns.canoliae
C; Genetics:
A;Gene: CC1512

10; 62; Gaps Query Match
7.5%; Score 101; DB 2; Length 216;
Best Local Similarity 22.7%; Pred. No. 0.18;
Matches 50; Conservative 29; Mismatches 79; Indels

78 QEVVRYQAYLQSRLGNYLP-PMSQLLTTARSWQACGHEPYQLPPEHLWGQIVPTLHLYQD 136 82 LEAVYWEN-----GDYVPDAVSALDKVLRDYRNDEVHPIDRGLYDLLDQIA-----RK 129 28 PQVSPIKTPSVLIT-----KDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATK 77 39 POATPVPPPTTVAATVASIDPPALKPAVDPRWVHLHN-------VHTGEK 81 Ор δλ g

137 LKSRGILPANTQIRSVYRNPELNQCAGG-----AAMSKHLTNSAIDIWVPDLEIKSQALY 191 δy qq

192 ELQNRLCQYWLEHGENQNFGLGLYAT-GAIHLDTQGFRKW 230 186 ALD-----LSVG-----GVGYYPTSNFVHVDVGPVRKW 213 Qγ pp

#### RESULT

Conserved hypothetical protein Atu3763 [imported] - Agrobacterium tumefaciens (strain C)Species: Agrobacterium tumefaciens (c)Species: Agrobacterium tumefaciens (species: Agrobacterium tumefaciens (c)Species: Agrobacterium tumefaciens (c)Species: O.Species: O.Spec AG3019

```
cyclable exported protein STY1255 [imported] - Salmonella enterica subsp. enterica se C; Species: Salmonella enterica subsp. enterica serovar Typhi A;Note: this species has also been called Salmonella typhi C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 09-Nov-2001 C;Accession: AH0044 R;Parkhill, J; Dougan, G; James, K.D; Thomson, N.R.; Pickard, D.; Wain, J.; Church H, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr, S.; Moule, S.; O'Gacra, P. Nature 413, 848-852, 2001 A;Atherford, R.; Simmonds, M.; Skelton, J.; Stevens, A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se A;Reference number: AB0502; PMID:11677608
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AH3310
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Scl. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
                                                                                                                                                                                                                                                                                            A; Molecule type: DNA
A; Residues: 1-637 <KUR>
A; Residues: 3 -637 <KUR>
A; Cross-references: GB:AE008918; PIDN:AAL53251.1; PID:g17984130; GSPDB:GN00191
A; Experimental source: strain 16M
C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A;Cross-references: GB:AL513382; PIDN:CAD08339.1; PID:g16502384; GSPDB:GN00176 C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --- GGAAMSKHLTNSAI 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        68 TWLQMHHATKQEVVRYQAYLQSRLG--NYLPPMSQLLTTARSWQACGHEPYQLPPEHLWG 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            86 ----QVYQSTGSREYI----TVVSAYRSPATNAMLRSSTRGVAKKSQHMLGRAMDYFIPG 137
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   72 MHHATKQEVVRYQAYLQSRLGNYLPP-MSQLLTTARSWQACGHEPYQLPP---EHLWGQI 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          128 VPTLHLYQDLKSRGILPANTQIRSVYRNPELN----QCAGGAAMSKHLTNSAIDIWVPD 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       51; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9 ITTLISSMLVACSAPIPTNPQVSPIKTP-SVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              37 VHTGEKAEI----AFKKD--GRFLPDGLKRLNVFLRDWRR--NEPTRMDPRLFDLIW--- 85
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          183 LEI-KSQAL---YELQNRLCQYWLEHGENQNFGLGLYATGA---IHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  138 VPLAKLRAIGMRYQIG------GVGYYPRSGSPFVHMDVGNVRHW 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7.1%; Score 96.5; DB 2; Length 637; 22.9%; Pred. No. 1.8; Live 31; Mismatches 53; Indels 5:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              6.9%; Score 93; DB 2; Length 263; 19.8%; Pred. No. 1.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               92; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            126 QIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            37; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 19.8%
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 40; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Molecule type: DNA
A; Residues: 1-263 <PAR>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Status: preliminary
                                                                                                                                                                                                                                                                                      A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A; Accession: AH0644
                                                                                                                                                                                                                                                       A; Accession: AH3510
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Map position: II
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Gene: BMEII0010
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Gene: STY1255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          В
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         α
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tunt A;Reference number: A97359; PMID:11743194
A;Reference number: A97359; PMID:11743194
A;Accession: B98265
A;Accession: B98265
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-634 < KUR>
A;Cross-references: GB:AE007870; PIDN:AAK89644.1; PID:g15159543; GSFDB:GN00170
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AH3510
hypothetical membrane associated protein BMEII0010 [imported] - Brucella melitensis (str
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hypothetical protein AGR_L_2143 [imported] - Agrobacterium tumefaciens (strain C58, Cere C; Species: Agrobacterium tumefaciens
C; Species: Agrobacterium tumefaciens
C; Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002
C; Accession: B98265
R; Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
                     A; Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
                                              ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; PMID:11743193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ٠.
ص
                                                                                                                                                                                                                                                   A; Cross_references: GB: AE008689; PIDN: AAL44573.1; PID: g17742190; GSPDB: GN00187
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              117 LFDLVWEVYRRSGATDYINVVSAFRSPETNGLLRTRTKGVAEKSQHMLGKAMDFYIPG-- 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             75 HTREKAVITFK-----RNGKYDQKGLQELNRFL---RDWRR--NQPTRMDPR----- 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      130 TLHLYQDLKSRGILPANTQIRSVYRNPELN----QCAGGAAMSKHLTNSAIDIWVPDLE 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     74 HATKQEVVRYQAYLQSRLGNY - - - - LPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIVP 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            74 HATKQEVVRYQAYLQSRLGNY----LPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIVP 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               130 TIHLYQDLKSRGILPANTQIRSVYRNPELN----QCAGGAAMSKHLTNSAIDIWVPDLE 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    28 HTREKAVITFK-----RNGKYDQKGLQELNRFL---RDWRR--NQPTRMDPR------ 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7.5%; Score 101; DB 2; Length 634; 22.5%; Pred. No. 0.71; tive 31; Mismatches 58; Indels 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  7.5%; Score 101; DB 2; Length 587; 22.5%; Pred. No. 0.64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            185 IKSQALYELQNRLCQYWLEHGENQNFGLGLYATGA---IHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            185 IKSQALYELQNRLCQYWLEHGENQNFGLGLYATGA---IHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              128 VKLATLREIGMOM------QIGGVGFYPTSGSPFVHMDVGGVRAW 166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   58; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   31; Mismatches
                                                                                                                                                                                                                                                                                            A; Experimental source: strain C58 (Dupont)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Map position: linear chromosome
                                                                                                                                                                                                                                                                                                                                                                                    A; Map position: linear chromosome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity 22.59
Matches 38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity
                                                                                                                                                                                                                                 A; Residues: 1-587 <KUR>
                                                                                                                                                                           A; Status: preliminary A; Molecule type: DNA
                                                                                                                                            A; Accession: AG3019
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Gene: AGR_L_2143
                                                                                                                                                                                                                                                                                                                                                       A;Gene: Atu3763
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                              C; Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        B98265
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            óλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
```

Gaps

12;

1 hypothetical protes space squence #status predi orotein #status predi 6.5%; Score 88;	Best_Loc Matches	OY   8 FITTLISSMIVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67   1   1   1   1   1   1   1   1   1	OY 126 QIVPTLHIYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAI 176  10	RESULT 8 Cy0815 hypothetical protein ECs1491 [imported] - Escherichia coli (strain 0157:H7 C;Species: Escherichia coli C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-200 C;Accession: C90815	R.Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, DNA Res. 8, 11-22, 2001. A.Title: Complete genome sequence of enterohemorrhagic Escherichia coli Ol A.Feference number: A99629; MUID:21156231; PMID:11258796 A.Stcession: C90815 A.Status: preliminary	A; Molecule type: DNA A; Molecule type: DNA A; Residues: 1-30 C.RAY> A; Residues: 1-30 C.RAY> A; Cross-references: GB:BA000007; PIDN:BAB34914.1; PID:g13360955; GSPDB:GNO A; Experimental source: strain 0157:H7, substrain RIMD 0509952 C; Genetics: A; Gene: ECs1491 C; Superfamily: conserved hypothetical protein b0819	Query Match Best Local Similarity 20.8%; Score 88; DB 2; Length 320; Best Local Similarity 20.8%; Pred. No. 4.1; Matches 53; Conservative 30; Mismatches 88; Indels 84; Gaps Oy 8 FITTLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67 ::
INLAE LYELQ  :   RYKAQ GFRKW	⋧	RESULT 6 #102215  two-component hybrid sensor and regulator all3275 [imported] - Anabaena sp. (strain PCC C. Species: Anabaena sp. A.Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120 C. bate: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 11-Jan-2002 C. Accession: AD2215	R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S DNA Res. 8, 205-213, 2001 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana A;Reference number: AB1807; MUID:21595285; PMID:11759840 A;Accession: AD2215 A;Status: preliminary A;Molecule type: DNA A;Residues: 1-889 kUR> A;Residues: 1-889 kUR> A;Cross-references: GB:BA000019; PIDN:BAB74974.1; PID:g17132370; GSPDB:GN00179 C;Genetics:	58; Ga	т <b>т</b>	QY 114 EPYQLPPEHLWQQIVPTHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTN 173	FESULT 7 F64855 YefS protein precursor - Escherichia coli C;Species: Escherichia coli C;Species: Escherichia coli C;Species: Escherichia coli C;Species: Escherichia coli C;Accession: F64855 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC R; Blattner, F64855 A;Title: The complete genome sequence of Escherichia coli K-12. A;Recession: F64855 A;Reference number: A64720; MUID:97426617 A;Reference number: A64720; MUID:97426617 A;Status: nucleic acid sequence not shown; translation not shown A;Residues: 1-320 cBLAT> A;Cross-references: GB:AE000211; GB:U00096; NID:g1787345; PIDN:AAC74197.1; PID:g1787356; C;Genetics: A;Gene: ycfS A;Gene: ycfS

```
nterohemorrhagic Escherichia coli 0157:H7 and g
231; PMID:11258796
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
; Shiba, T.; Hattori, M.; Shinagawa, H.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1] - Escherichia coli (strain 0157:H7, substrain
                                                                                                                                                          11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             e 88; DB 2; Length 320;
1. No. 4.1;
dismatches 88; Indels 84; Gaps 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           BAB34914.1; PID:913360955; GSPDB:GN00154 substrain RIMD 0509952
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18-Jul-2001 #text_change 03-Aug-2001
                                                                                                                                                                                                                                                                                                                                              RNPELNQCA-----GGAAMSKHLTNSAI 176
                                                                                                                                                                                                                                                                                                                                                                                                                         ------RLCOY---WLEHGENO 208
                                                                                                                                                                                                                                                                       YLPPMSQLLTTARSWQACGHEPYQLPPEHLWG 125
                                                                                                                                                                                                                                                                                                                                                                          YPPGKNSVTVYPIGIGQLGGDTLTPTMVTTVS 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                        AVVPAGLDNPMGHHAIRLAAYGGVYLLHGTNA 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /LPPMSQLLTTARSWQACGHEPYQLPPEHLWG 125
                                                                                                                                                          84; Gaps
                                                                                                                                                                                                                                                                                             |:| :||
XVPRAGSVLT------89
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RNPELNQCA-----GGAAMSKHLTNSAI 176
                                                                                                                                                                                        PSVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67
                                                                                                                                                                                                                  PSVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |:| :||
XVPRAGSVLT------ 89
                                                                                                               re 88; DB 2; Length 320;
1. No. 4.1;
4ismatches 88; Indels
                                      predicted <SIG>
predicted <MAT>
protein b0819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein b0819
```

Wed Jul 31 08:19:59 2002

```
phosphoribosylformylglycinamidine synthetase I [imported] - Listeria monocytogenes (S C; Species: Listeria monocytogenes C; Species: Listeria monocytogenes C; Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001 C; Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001 C; Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #sequence_revision 27-Nov-2001 E; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec C; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, D.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001 A; Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ox, C.; Schlueter, T.; Sinces, N.; Tierera, A.; Vazquez-Boland, J.A.; Voss, H.; Wehla A; Title: Comparative genomics of Listeria species.
A; Reference number: AB1077; MUID:21537279; PMID:11679669 A; Status: preliminary A; Molecule type: DNA A; Residues: 1-739 <GLA> A; Cross-references: GB:NC_003210; PIDN:CAC99847.1; PID:g16411223; GSPDB:GN00177 A; Experimental source: strain EGD-e
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            13;
                      11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    :: |:: |:: | 407 DVVGV---WKELLAQPIIASKRHIYEQ-----YDYQVRTDTAVVPGSDAAIVRVRGTEKA 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      352 VIDDKM--YKIIHHGEVVANVPVDALAED-APVYHKPSKEPTRYQAFQEEA--FVPAMD 406
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   145 ----ANTQIRSVYRNPELN------QCAGGAAMSKHLTNSAIDIWVPDLEIKSQ 188
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  189 ALYELQ---NRLCQYWLE-----HGE----NONFGLGLYATGAI------HLDTQ 225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        100 QLLTTARSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILP--------- 144
                                                                                                                                                                                                                                                                83 YQAYLQSRLGNYLPPMSQLL-TTARSWQACGHE-----PYQLPPEHLWGQIVPTLHLYQ 135
                                                                                                                                                                                                                                                                                           136 DLKSRGI-LPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAID-----IWVPDL---EI 185
                                                                                                                                                                                                                                                                                                                                                                                                          927 LLQNQAIPLPSAQAWEYIKKTDSASQCI-----KKLLDSEEDPLTSSFWWFHNRDTEEI 980
                                                                                                                   754 LLSSKTPLPTTNYLSAFTESLYTDVQDSVSKRLETLQKDPATTPFSVVFSDQLFHDPLHP 813
                                                                                                                                                                                                                814 NDQQVSPLLLSLPLKNIHLQGTIHGVCSKGVYLFSMHPG-EAFKKT--QKTHGFPKDAFE 870
                         84; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            40 ITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMS 99
                                                                      ----PIKTPSVLITKDKIGDHHTHE 52
                                                                                                                                                                -----HVGLQAHFETWLQMHHATKQEVVR 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        II
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gene: purQ Superfamily: phosphoribosylformylglycinamidine synthase component
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              73;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 739;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              83; Indels
                         88; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 42; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match 6.5%; Score 87.5; I Best Local Similarity 19.5%; Pred. No. 13; Matches 48; Conservative 42; Mismatches
red. No. 17;
Mismatches
     Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       981 CSRLSNDVLSQLLSLFINQDSQON 1004
                                                                                                                                                                                                                                                                                                                                                                                                                                                            186 KSQALYELQNRLCQYWLEHGENQN 209
                         40;
                                                                                                                                                                        53 HDESVS------
19.78;
  Local Similarity 19.79
nes 52; Conservative
                                                                           17 LVACSAPIPTNPQVS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            573 DFKNSG 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  226 GFRKWG 231
       Best Loca
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  οy
                                                                                                                                                                                                                                                                                                                                                                                                               Dβ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              QΥ
                                                                           οy
                                                                                                                        g
                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                  δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                        δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A;Residues: 1-1004 <ARN>
A;Cross-references: GB:AE001334; GB:AE001273; NID:g3329078; PIDN:AAC68244.1; PID:g332908
A;Experimental source: Serotype D, strain UW-3/CX
C;Genetics:
                                                                                                                                                              hypothetical protein ycfS [imported] - Escherichia coli (strain O157:H7, substrain EDL93 C; Species: Escherichia coli C; Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001 C; Accession: G85674 R; Perion W.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhewiller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, Nature 409, 529-533, 2001
                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Molecule type: DNA
A;Residues: 1-320 <STO>
A;Cress-references: GB:AE005174; NID:912514663; PIDN:AAG55859.1; GSPDB:GN00145; UWGP:217
A;Experimental source: strain 0157:H7, substrain EDL933
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          126 QIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCA-----GGAAMSKHLTNSAI 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                68 TWLQMHHATKQEVVRYQAYLQSRLG--NYLPPMSQLLTTARSWQACGHEPYQLPPEHLWG 125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             148 DKRANPTWTPTANIR--ARYKAQGIELPAVVPAGPDNPMGHHAIRLAAYGGVYLLHGTNA 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      8 FITTLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --- RLCQY ---WLEHGENQ
                                                                                                                                                                                                                                                                                                                              A;Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7. A;Reference number: A85480; MUID:21074935; PMID:11206551 A;Accession: G85674
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.5%; Score 88; DB 2; Length 1004;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      6.5%; Score 88; DB 2; Length 320; 20.8%; Pred. No. 4.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ---SLEAIAKKYNVGFLALLQANPGVDPYVPRAGSVLT---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Gene: ycfs
C;Superfamily: conserved hypothetical protein b0819
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            177 D----IWVPDLEIKSQALYELQN----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match 6.5%
Best Local Similarity 20.8%
Matches 53; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       :||:|: ::| | |
206 DFGIGMRVSSGCIRL 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           209 NFGLGL-YATGAIHL 222
          209 NFGLGL-YATGAIHL 222
                                     :||:|: ::| | |
206 DFGIGMRVSSGCIRL 220
                                                                                                                                                                                                                                                                                                                                                                                                                      A; Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Gene: recC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          QΥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            d
            δ
```

RESULT 12	Db 128 LLVAPLAGISINPPGSKFTFATNVYTPQALGYSRDSDSDPA168
A70523 Probable oxidoreductase - Mycobacterium tuberculosis (strain H37RV) C;Species: Mycobacterium tuberculosis C;Date: 177u1-1998 #sequence revision 177u1-1998 #fewt Ahange 22-AA+-1000	OY 76 TKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACG-HEPYQLPPEHLW 124  Db 169RYQGREVSLQRLTYFSPTLAYQVNDELSVGFSHQAVALNEDFRAPNQ 219
C; Accession: A70523  R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. Comor, R.; Davies, R.; Devlin, K.; Peltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S. Rajandream, M.A.; Roqers, J.; Rutter, S.; Seeder, K. ekalton, C. comoros, C. Comoro	QY 125 GQIVPTLHLYQDLKSRGILPANTQIRSVTRNPELNQCAGGAAMSKHLTNSAID 177  DD 220LIGLLQQTKEIGCLPGMQEILEVFFNVCGGNIGPFQNLANIDLDMQQSLSP 270
Nature 393, 537-544, 1998 A; Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A; Reference number: A70500; MUID:98295987	OY 178GAI 220 
A) Status: preliminary; nucleic acid sequence not shown; translation not shown b) Modecule type; DNA	RESULT 14 E36792
A; residues: 1-350 <col/> A; residues: 1-350 <col/> A; residues: 1-350 <col/> A; Experimental source: GB:297188; GB:AL123456; NID:g3261805; PIDN:CAB10023.1; PID:e1300070; A; Experimental source: strain H37Rv C; Genetics: A; Gene: RV3829c	hypothetical protein ORF58 - ictalurid herpesvirus 1 (strain auburn 1) C;Species: ictalurid herpesvirus 1 A;Note: host Ictalurus punctatus (channel catfish) C;Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 24-Nov-1999 C;Accession: R36792
Query Match 6.3%; Score 86; DB 2; Length 536; Best Local Similarity 24.0%; Pred. No. 12; Matches 44; Conservative 22; Mismatches 49; Indels 68; Gaps 11;	R; Davison, A.J. submitted to GenBank, January 1992 A; Description: Channel catfish virus: a new type of herpesvirus. A; Reference number: A36804 A; Accession: P36702
IDESVSHVGLQAHFETW 69 :     ::   ::   .:	A; Molecule type: DNA A: Residues: 1-599 < DAV> A; Cross references: GB:M75136; NID:9331209; PIDN: AAA88161.1; PID:9331268 R: Davison A J
PYQL	Virology 186, 9-14, 1992 A; Title: Channel catfish virus: a new type of herpesvirus. A; Reference number: A39447; MUID:92087490 A; Contents anotation
SKHLTNSAIDI 1	her protein nor nucleic acid sequence is given
* Ed. 30	C;Superiamily: ictalurid herpesvirus 1 hypothetical protein ORF58
Db 406 WFP 408	Query Match 6.2%; Score 84; DB 2; Length 599; Best Local Similarity 18.2%; Pred. No. 20; Matches 58; Conservative 50; Mismatches 95; Indels 116; Gaps 15;
RESULT 13 83424 hypothetical protein PA1764 [imported] - Pseudomonas aeruginosa (strain PAO1) C;Species: Pseudomonas aeruginosa	QY 1 MKNFNQYFITTLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTH 51
C; Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000 C; Accession: G83424 R; Stover: C.K.; Phan, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Bradaman, S. Yman, V. Prodi, T. T. Control	QY 52 EHDESVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQAC 111 :::::  :  :  :  :  :  Db 327 DNSKFLAHVVMDRAMPTTCYIHDDIDVKAMIOSYISRVYRMIODTOVRICELSAA 381
Larbıg, ortunist	
A; Accession: G83424 A; Status: preliminary A; Molecule type: DNA A; Residues: 1-512 < STO	OY 142 ILPA 167  S   1   1   1   1   1   1   1   1   1
A;Cross-references: GB:AE004602; CB:AE004091; NID:g9947739; PIDN:AAG05153.1; GSPDB:GN001 Experimental source: strain PAO1 C;Genetics: A;Gene: PA1764	
Length 532;	OY 196 RLCOYWLEHGENONF 210                         Db 560 THCAGCADFWLTHGSDPNF 578
MLVACS	RESULT 15 B82162 hypothetical protein VC1750 [imported] - Vibrio cholerae (strain N16961 serogroup 01)

```
A.Molecule type: DNA
A.Residues: 1-373 <HEL>
A.Cross-references: CB:AE004252; GB:AE003852; NID:g9656263; PIDN:AAF94900.1; GSPDB:GN001
A.Experimental source: serogroup 01; strain N16961; biotype El Tor
C.Genetics:
                                                       Ryleidelberg, J.E.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Selhers, I.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833
A;Accession: B82162
A;Status: preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                14;
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: B82162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75 ATKQEVVRYQAYLQSRLGNY-----LPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIV 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -----QDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNS 174
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               173 QWFGLTYVSGFHEQASAFSPEDLYS-NMLGANLARDVLLANPDANKQEFEKIFAHLLEDE 231
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ouery Match 6.2%; Score 83.5; DB 2; Length 373;
Best Local Similarity 23.3%; Pred. No. 12;
Matches 64; Conservative 32; Mismatches 88; Indels 91; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               8 FITTLISSMLVACSAP---IPTNP-------QVSPIKTP--SV--LITKDKI 45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 FITTLL-CFLLAILQPVQAIPTPPIGLRPCCAFGYNLHAQVAGIPVPFFSVDNVIDVDAL 61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         46 GDHHTHEHDESVS-----QMHH 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       175 AIDIWVPDLEIKSQALYEL-----QNRLCQYWL 202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        129 PTLHLY----
                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Map position: 1
                                                                                                                                                                                                                                                                                                                                                                                                                       A; Gene: VC1750
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ρp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           δŽ
```

Search completed: July 30, 2002, 16:15:54 Job time: 447 sec

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

July 30, 2002, 16:15:13 ; Search time 11.07 Seconds (without alignments) 874.425 Million cell updates/sec Run on:

US-09-674-779-2 1355 1 MKNPNQYFITTLISSMLVAC.......GAQFSETNSICRHVLPKNKL 250 Title: Perfect score: Sequence:

BLOSUM62 Gapop 10.0 , Gapext 0.5 Scoring table:

105224 seqs, 38719550 residues Searched:

105224 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution. SwissProt\_40:\*

#### SUMMARIES

	Description	P75954 escherichia	Q00157 ictalurid h			09kd49 bacillus ha			_				arabidc						4	085045 listeria mo	~			Q9kgu2 thermoanaer	P06516 synechococc	-	073700 gallus gall					m	$^{1}$ ps	. 7
SUMMAKIES	ID	YCFS_ECOLI	VG58_HSVI1	LOLB_HAEIN	YAC8_YEAST	SYGA_BACHD	MLTA_ECOLI	PEDA_LACHE	CD5S_MOUSE	NIR_SPIOL	REDD_STRCO	VL1_BPV4	ALA2_ARATH	COAA_BACHD	RBMA_RAT	MUTS_BACHD	VASS_BPGA	PHNL_DESFR	CATO_HUMAN	ALR_LISMO	YUY1_CAEEL	VNSS_TSWVL	KPCM_HUMAN	XYLA_THEYO	CAPP_SYNP6	PINE_MOUSE	CCAD_CHICK	RRPL_EBOSM	TRB2_AERPE	YE72_HUMAN	CPO_DROME	- 1	- 1	XYLA_BACLI
	Length DB	320	599 1						369 1	594 1	350 1	506 1	1107 1		85,2 1	865 1			321 1													1440 1		
\$ Ouerv	. 1	5.	•	•	•	٠		٥.	σ.	٠.۲	· ·	۲.		. و	9.	9.	. ب	9.	ı.	٠.	ı.	٠. د	ų.	ΰr	٥٠١	ָי.	יי	٠.		٠.	5.4	4.4	5.4	5.4
	Score	88	84	83	83	82.5	82.5	80.5		77.5	//	//	1	76.5	76.5	- 1	75.5	ഹ	75	2,1		74.5	₹"	4,1	4 .	/ 4	4,	٦,	73.5	÷.	m .	m	73	
Result	No.	-	2 (	۰ ٦٠	41	n '	ا ف	,	ж (	ກຸ	o F	17	77	5 T	T 4	15	91	17	80 0	5 6	200	77	77	5 7 6	7 0	0.7	9 10	77	87	29	30	31	3.2	33

P49803 rattus norv P49802 homo sapien	P10394 drosophila	P20806 drosophila	Q9nrco nomo sapien Q81023 human papil	P26191 porcine rot	084417 chlamydia t P28955 equine herp	P53068 saccharomyc	Q59940 streptococc
RGS7_RAT RGS7_HUMAN	POL4_DROME	7LES_DROVI	VL2_HPV54	VP2_ROTPC	FMFA_CHLTR TEGU_HSVEB	YGZ0_YEAST	IDH_STRMU
		<del></del>				Η,	-
477	1237 1273	2594	470	872	3421	283	393
5.4	5.4	7. r	5.4	о 4.	. 4.	 	5.3
73	73	73	72.5	72.5	72.5	72	7/
34 35	36 37	88 G	40	41	4.3	44	40

#### ALIGNMENTS

```
STANDARD;
                                                                                                                                        168 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=727;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    influenzae Rd.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TIGR; HI1607
                                                                                                                                                                                                                                                                                                                                                                                                   Haemophilus
                                                                                                                                                                                                                                                                                LOLB_HAEIN
P45270;
                                                                                                                                                                                                                                                                   LOLB_HAEIN
                                                                                                                                                                                                                                                                                           δλ
                                                                                     Óλ
                                                                                                                g
                                                                                                                                         ŏ
                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                           11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      267 MAMFRRYPIDLRLERMIMDYFGSGLKCTLATFTPPKYTVSGERSRHSIVIKNHVDKHYLN 326
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 52 EHDESVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQAC 111
                                                                                                                                                                                                      ---GGAAMSKHLTNSAI 176
                                                                                                                                                                                                                                                                        68 TWLQMHHATKQEVVRYQAYLQSRLG--NYLPPMSQLLTTARSWQACGHEPYQLPPEHLWG 125
                                                                                                                                                                                                                              90 --IPLQTLLPDAPREGIVINIAELRLYYYPPGKNSVTVYPIGIGQLGGDTLTPTMVTTVS 147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       95; Indels 116; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1 MKNFNQYFITTLISSML-----VACSAPIPTNPQ--VSPIKTPSVLITKDKIGDHHTH 51
                                                                           Gaps
                                                                                                  8 FITTLISSMLVACSAPIPINPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFE 67
                                                                                                                                                                                                                                                        177 D----IWVPDLEIKSQALYELQN--------RLCQY---WLEHGENQ
                                                                                                                          10 WLTFFTFAAAVALALPAKANTWPLP-PAGSRLVGENKF---HVVENDGG-----
                                                                           84;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6.2%; Score 84; DB 1; Length 599; 18.2%; Pred. No. 4.9; tive 50; Mismatches 95; Indels
                                                Query Match 6.5%; Score 88; DB 1; Length 320; Best Local Similarity 20.8%; Pred. No. 0.99;
                                                                           88; Indels
320 HYPOTHETICAL PROTEIN YCFS.
34636 MW; 89378A8DFD60359B CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       599 AA; 67476 MW; 991C2B524F0F6BEF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      "Channel catfish virus: a new type of herpesvirus.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
unclassified Herpesviridae.
                                                                                                                                                                           ---SLEAIAKKYNVGFLALLQANPGVDPYVPRAGSVLT----
                                                                                                                                                                                                       126 QIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCA----
                                                                                                                                                                                                                                                                                                                                                                                                                       COLLON-1992 (Rel. 24, Created)
01-DEC-1992 (Rel. 24, Last Sequence update)
01-DEC-1992 (Rel. 24, Last annotation update)
                                                                           30; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=92087490; PubMed=1727613;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-DEC-1992 (Rel. 24, Last and
Hypothetical gene 58 protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; M75136; AAA88161.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best Local Similarity 18.29
Matches 58; Conservative
                                                                            53; Conservative
                                                                                                                                                                                                                                                                                                           209 NFGLGL-YATGAIHL 222
                                                                                                                                                                                                                                                                                                                                   DFGIGMRVSSGCIRL 220
                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Virology 186:9-14(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            protein
 24 3
320 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. STRAIN-AUBURN 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=10401;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Davison A.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hypothetical
                                                                                                                                                                                                                                                                                                                                                                                    VG58_HSVI1
ID VG58_HSVI1
AC Q00157;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Qγ
                                                                                                                                                                                                                                 QQ
                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                 Db
                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                FT
                                                                                                                              g
                                                                                                                                                        δŏ
                                                                                                                                                                                g
                                                                                                                                                                                                       ò
```

```
between the Swiss Institute of Bioinformatics and the EMBL outstation—the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
Chaperone; Outer membrane; Lipoprotein; Transport; Protein transport;
Signal; Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       502 YIATAMLHILHGENLETHDLGPQTGTGVVGEPPKHCTGAQLREWFIDLQ-KTLAL-DAPS 559
::::||::|
327 DNSKFLAHVVMDRAMPTTCYIH-----DDIDVKAMIQSVISRVVRMIQDTQVRLQELSAA 381
                                                                                                                                                                                                                                                                                                                                                                                                                                                          -----SKHLTNSAIDIWVPDLEIKSQALYELQN 195
                                                                                                                   -----HLWGQI-VPTL-----HLYQDL--KSRG 141
                                                                                                                                                                                                  382 GNKLFHMFFNQLPPEMNNLDISFKYNPLVDHALQGQKGVPGLQYGNPEQIYTDMITEMTA 441
                                                                                                                                                                                                                                                                                                                                                                      442 LLPRIGHMVEFMEWSQELSSSKTLCLLLPKIARDLNVPPLVTSLNDDTGNAMLIAFAHKI 501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Relaxed A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., Mcravage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -1-FUNCTION: PLAYS A CRITICAL ROLE IN THE INCORPORATION OF LIPOPROFIENT IN THE OUTER MEMBRANE AFTER THEY ARE RELEASED BY THE LOLA PROPEIN (BY SIMILARITY).
-1- SUBCELLUIAR LOCATION: Attached to the outer membrane by a lipid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         BY SIMILARITY.
OUTER-MEMBRANE LIPOPROTEIN LOLB.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 anchor (By similarity).
-!- SIMILARITY: BELONGS TO THE LOLB FAMILY.
-!- CAUTION: WAS ORIGINALLY THOUGHT TO BE INVOLVED IN DELTA-
                                                                                                                                                                                                                                                                                           142 ILPA-----PELNQCAGGAAM--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Whole-genome random sequencing and assembly of Haemophilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
16-CCT-2001 (Rel. 40, Last annotation update)
Outer-membrane lipoprotein lolB precursor.
LOLB OR HERM OR HI1607.
Haemophilus influenzae.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              209 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AMINOLEVULINIC ACID BIOSYNTHESIS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-RD / KW20 / ATCC 51907;
MEDLINE-95350630; PubMed-7542800;
                                                                                                                                112 GHEPY----QLPPE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       196 RLC----QYWLEHGENQNF 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   560 THCAGCADFWLTHGSDPNF 578
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; U32834; AAC23251.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Science 269:496-512(1995).
```

```
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                           59 HVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGHEPYQL 118
                                                                                                                                                                                                                                                        119 PPEHLWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDI 178
                                                                                                                                                                                                                                                                                                                     179 WVPDLEIKSQALYELQNRLCQYWL--EHGENQNFGLGL-YATGAIHLDTQGFRKWGAQ-- 233
                                                                                                                                                                                                                                                                                       81 PKSY-----TLKLYSLISKSTLWIQMHQSGMTISDNNGNQ--QSAANSKLLLQEIIGM 131
                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                             1 MKNFNQYFITTLISSMLVACSAPI--PTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVS 58
                                                                                                                                                                                                               42 ------W-QQHLQKIQXQA--KGQIG-YISP-TERFSSRFEWQ-----YQN 80
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=S288C / AB972;
MEDLINE=95249563; PubMed=7731988;
Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.,
Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,
 22 N-ACYL DIGLYCERIDE (BY SIMILARITY).
24193 MW; 0555F02F13E852A1 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
                                                                                                      74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   YAC8_YEAST STANDARD, PRT; 528 AA.
P3973; P3973;
Ol-FEB-1995 (Rel. 31, Created)
Ol-OCT-1996 (Rel. 34, Last sequence update)
Ol-OCT-1996 (Rel. 34, Last annotation update)
Hypothetical 58.8 kDa protein in MYO4-DRS2 intergenic region.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              nucleotide sequence of chromosome I from Saccharomyces
                                                                6.1%; Score 83; DB 1; Length 209; 23.0%; Pred No. 1.7;
                                                                                                  87; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         E5A3CC7C6D60977A CRC64;
                                                                                                                                                Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995)
                                                                         Saccharomyces cerevisiae (Baker's yeast).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-SER. POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein; Transmembrane.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           58750 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; U12980; AAC05004.1; -.
                                                                                               59; Conservative
                                                                                                                                                                                                                                                                                                                                                                                        234 -FSETNSICRHVLPKN 248
                                                                                                                                                                                                                                                                                                                                                                                                                     178 TYHSNNSMPENILLKN 193
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          S0000026; YAL028W.
               209 AA;
                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         528 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NCBI_TaxID=4932;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cerevisiae.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Hypothetical
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REVISIONS.
                 SEQUENCE
                                                                Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "The
FT
                                                                                                                             à
                                                                                                                                                                                            δ
                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                          δλ
                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                     Qγ
                                                                                                                                                                                                                                                                                                                                                                                      δ
```

```
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                   16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
61ycyl-tRNA synthetase alpha chain (EC 6.1.1.14) (Glycine--tRNA ligase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      63 QAHFETWLQMHHATKQEVVRYQAYLQSRLGNY---LPPMSQLLTTARSWQACGHEPYQL- 118
                                                                                                                                               368 QLQ-QTEALLKHSLKDEVLKDENDLVKNIANFDKIVKELRDLRSRTIGWKELVEEDYLMN 426
                                                                                                                                                                                    119 ------PPEHLWGQIVPTLHL----YQDLKSRGILPANTQIRSVYRNPELNQCAGGAA 166
                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -i- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PROSITE; PS00179; AA_TRNA_LIGASE_II_1; FALSE_NEG.
PROSITE; PS00339; AA_TRNA_LIGASE_II_2; 1.
Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
                                                             3 NFNQYFITTLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGL 62
                                     46;
   6.1%; Score 83; DB 1; Length 528;
19.4%; Pred. No. 5.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            + 91ycyl-trna(Gly).
-!- SUBDNIT: TETRAMER OF TWO ALPHA AND TWO BETA CHAINS (BY SIMILARITY).
                                   89; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            297 AA; 34420 MW; F277D6C2C74969E7 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                  427 LKODFDKENPESFEARLSDTINTNVAKLQDLEKR----
                                35; Mismatches
                                                                                                                                                                                                                                                                              470 SRKEVMR------- 488
                                                                                                                                                                                                                                               167 MSKHLTNSAIDIWVPDLEIKSQALYELQNRL 197
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro; IPR002106; AA_tRNA_ligase_II.
InterPro; IPR002310; tRNA_synt_2e.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -!- SUBCELLULAR LOCATION: Cytoplasmic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam, PF02091; tRNA_synt_2e; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AP001511; BAB05089.1; -
              Best Local Similarity 19.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRINTS; PR01044; TRNASYNTHGA
                                41; Conservative
                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   alpha chain) (GlyRS).
GLYQ OR BH1370.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bacillus halodurans.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Completé proteomé.
SEOUENCE 297 AA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NCB1_TaxID=86665;
                                                                                                                                                                                                                                                                                                                                                           SYGA_BACHD
Query Match
                                                                                                                                                                                                                                                                                                                                                                        Q9KD49;
                                                                                                                                                                                                                                                                                                                                         SYGA_BACHD
                                Matches
                                                                                         Dp
                                                                                                                        δλ
                                                                                                                                                  g
                                                                                                                                                                                                                a
                                                                                                                                                                                                                                                                                                                                                                      ŏ
                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                           g
```

```
12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN-0157:H7 / RIMD 0509952;
MEDIINE-21156231; PubMed=11258796;
Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
Hayashi T., Makino K., Ohnishi M., Kurokawa T., Tanaka M., Tobe T.,
Itan C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MLTA_ECOLI STANDARD; PRT; 365 AA.
P468B5; P76638; Q4662B;
01-N0V-1995 (Rel. 32. Created)
01-N0V-1997 (Rel. 35. Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Membrane-bound lytic murein transglycosylase A precursor (EC 3.2.1.-)
(Murein hydrolase A) (M113B).
Escherichia coli, and
                                                                                                                                                                                                                                                                                                    125 ------LGWEVWLDGMEITQFTYFQQVGGLEANPVSAEITYGLE-RLASY-IQDKE 172
                                                                                                   67 ETWIQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTA-RSWQACGHEPYQLPPEHLWG 125
                                                                                                                                                                                    126 ------YRNPELNQCAG 163
                                                                                                                                                                                                            66 ENPNRLYQHHQFQVIMKPSPTHIQELYLDSLRALGINPLEHDIRFVEDNWENPSLG-CAG 124
                                                                                                                                                                                                                                                                           164 GAAMSKHLTNSAIDIWVPDLEI---------KSQALYELQNRLCQYWLEHGE 206
                                                          56; Indels 59; Gaps
                                                                                                                                             12 EYW-----SKONCILLIQAYDTEKGAGTMSPYTMLRTIGPEPWNVAYVEPSRRPADGRYG 65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE-97431497; PubMed=9287002;
MEDLINE-57431497; PubMed=9287002;
Lommatzsch J., Templin M.F., Kraft A.R., Vollmer W., Hoeltje J.-V.;
"Outer membrane localization of murein hydrolases: MltA, a third
lipoprotein lytic transglycosylase in Escherichia coli.";
J. Bacteriol. 179:5465-5470(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
MEDILINE-97426617; PubMed-9278503;
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-015:H7 / EDL933 / ATCC 700927;
MEDLINE-21074935; PubMed-11206551;
MEDLINE-21074935; PubMed-11206551;
Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
Rose D.J., Mayhew G.E., Evans P.S., Gregor J., Kirkpatrick H.A.,
Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
Welch R.A., Blattner F.R.;
"Genome sequence of enterchaemorrhagic Escherichia coli O157:H7.";
Nature 409:529-533(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
                  DB 1; Length 297;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               "The complete genome sequence of Escherichia coll K-12."; Science 277:1453-1474(1997).
Score 82.5; DE
                                    24.3%; Pred. No. 2.8; tive 22; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A., AND CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Escherichia coli 0157:H7.
                                                          44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NCBI_TaxID=562, 83334;
                                        Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                            173 N 173
                                                                                                                                                                                                                                                                                                                                                                 207 N 207
                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MLTA_ECOLI
                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   οy
                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                   ò
                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                            ò
                                                                                                                                                  g
                                                                                                                                                                                             QΥ
                                                                                                                                                                                                                                     셤
```

```
Wolling W., von Rechenberg M., Hoeltje J.-V.;

Wyolling W., von Rechenberg M., Hoeltje J.-V.;

The publis, the lytic transglycosylase Mita, and the scaffolding protein pells, the lytic transglycosylase Mita, and the scaffolding protein mipa of Escherichia coli ":

"Biol. Chem. 274:6726-6734(1999).

"Biol. Chem. 274:6726-6734(1999).

"PRINTENTON: MIRENI-DEGRADING ENZYRE. MAY PLAY A ROLE IN RECYCLING OF MUROPEPTIDES DURING CELL BONDATION AND/OR CELL DIVISION. OPTIMAL ACTIVITY IS BETWEEN PH 4.0 AND 4.5; LOSES ITS ACTIVITY RAPIDLY AT TEMPERATURES ABOVE 30 DEGREES CELSIUG. DEGRADES MUREIN GINCAN CATANATIC ACTIVITY: CLEAVAGE OF THE BETA-1, 4-GLYCOSIDIC BOND STRANDS AND INSOLUBLE, HIGH-MOLECULAR WEIGHT MUREIN SACCULI.

"CATALYTIC ACTIVITY: CLEAVAGE OF THE BETA-1, 4-GLYCOSIDIC BOND BETWEEN N-ACETYLAURANIC ACID AND N-ACETYLOLICOSAMINE RESIDUES.

"THEREBY CONSERVING THE ENERGY IN A NEMLY SYNTHESIZED
"THEREBY CONSERVING THE BURRAIL CALD RESIDUE."

"THEREBY CONSERVING THE MURANIC ACID RESIDUE."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its most by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GNYLPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQIRS 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      67 RNSSPRLYGNQSNVYNAVQEWLRAGGDTRNMRQFGIDAWQMEGADNYGNVQFTGY---- 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            6 QYFITTLISSMLVACSA-PIPTNPQVSPIK--TPSVLITK-DKIG-DHHTHEHDESVSHV 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -- OMHHATKQEVVRYQAYLQSRL 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             88; Indels 111;
                                             0157:H7 and genomic comparison with a laboratory strain K-12."; DNA Res. 8:11-22(2001).
                                                                                                                        SEQUENCE OF 147-161; 204-213 AND 258-280, AND CHARACTERIZATION.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.1%; Score 82.5; DB 1; Length 365; 19.1%; Pred. No. 3.7; Live 43; Mismatches 88; Indels 111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TRANSGLYCOSYLASE A.
N-ACYL DIGLYCERIDE (PROBABLE).
THE N - N (IN RRE 1).
SECBB92C1E8D5969 CRC64;
Kuhara S., Shiba T., Hattori M., Shinagawa H.; "Complete genome sequence of enterohemorrhagic Escherichia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEMBRANE-BOUND LYTIC MUREIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Cell wall; Hydrolase; Glycosidase; Signal; Lipoprotein; Outer membrane; Multigene family; Complete proteome. SIGNAL 1 20 PROBABLE.
                                                                                                                                                                Ursinus A., Hoeltje J.-V.;
"Purification and properties of a membrane-bound lytic transglycosylase from Escherichia coli.";
J. Bacteriol. 176:338-343(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ECOGENE; EG13085; mltA.
PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
                                                                                                                                                                                                                                                                                     INTERACTION WITH MIPA AND MRCB/PONB. STRAIN-ATCC 53338 / MC1061;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL, U29581, AAB40463.1, ALT_INIT.
EMBL, AE005509, AAG57927.1; -..
EMBL, AP002563; BAB37096.1; -.
                                                                                                                                                                                                                                                                                                                      STRAIN=ATCC 53338 / MC1061;
MEDLINE=99156961; PubMed=10037771;
                                                                                                                                                 MEDLINE-94117367; PubMed-8288527;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               21 21 N
346 346 H
365 AA; 40410 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61 -----GLQAH----FETWL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL; U32224; AAC45723.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EMBL; AE000365; AAC75855.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Local Similarity 19.1
Matches 57; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  365
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           anchor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LIPID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
            ò
```

16;

150 RSVYRNPELNQCAGGAAMSKHLTNSAIDIWVPDLEIKSQALYELQNRL 197

QΩ Qγ

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -!- CATALYTIC ACTIVITY: DIPEPTIDE + H(2)0 = 2 AMINO ACID.
-!- ENZYME REGULATION: INHIBITED BY ZN(2+), CU(2+), CA(2+) AND CD(2+).
                                  152 VYRNPELNQCAGGAAMSKHL---TNSAIDIWVPDLEIKSQALYELQNRLCQYWLEHGENQ 208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Dudley E.C., Husgen A.C., He W., Steele J.L.; "Sequencing, distribution, and inactivation of the dipeptidase A gene (pepDA) from Lactobacillus helveticus CNR232."; J. Bacteriol. 178:701-704(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Vesanto E., Peltoniemi K., Purtsi T., Steele J.L., Palva A.;
"Molecular characterization, over-expression and purification of a
novel dipeptidase from Lactobacillus helveticus.";
Appl. Microbiol. Biotechnol. 45:638-645(1996).
-I- FUNCTION: HYDROLYZES A WIDE RANGE OF DIPEPTIDES BUT UNABLE TO
HYDROLYZE DIPEPTIDES CONTAINING PROLINE. HIGHEST ACTIVITY AGAINST
       -----RGRLPSRAEIYA 158
                                                                             ---YIDFGDGS 193
                                                                                                        209 NFGLGLYA-----TGAIHLD-----TQGFRKWGAQFSETNSICRHVLPKN 248
                                                                                                                                      194 PLNFFSYAGKNGHAYRSIGKVLIDRGEVKKEDMSMQAIRHWGETHSEAE--VREILEGN 250
                                                                                                                                                                                                                                                                                                                                                                                Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
Lactobacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -i- ENZYME REGULATION: INHIBITED BY ZN(2+), CU(2+), CA(2+) AND CE-i- SUBUNIT: HOMOOCTAMER (PROBABLE).
-i- MISCELLANEOUS: OPTINAL ACTIVITY IS OBSERVED AT PH 6.0 AND 55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    T -> A (IN REF. 2).
3BFA79983D3ECEF2 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY U34.
                                                         159 ------GALSDKYILAYSNSLMDNFIMDVQGSG---
                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                       474 AA
122 --YTPVIQARHTRQGEFQ---YPIYRMPPK-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [2] SEQUENCE FROM N.A., AND CHARACTERIZATION.
                                                                                                                                                                                                               MEDLINE=96146518; PubMed=8550503;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=96338998; PubMed=8766699;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               474 AA; 53512 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; U34257; AAC43971.1; -. EMBL; Z38063; CAA86210.1; -.
                                                                                                                                                                                                                                                                                                                                                             Lactobacillus helveticus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hydrolase; Dipeptidase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DEGREES CELSIUS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                  NCBI_TaxID=1587;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEROPS; U34.001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=CNRZ32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                 PEDA_LACHE
                                                                                                                                                                                                                                  ò
                                                                    pp
                                                                                                        ŏ
                                                                                                                                          g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license gareement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                  MEDLINE-9832126; PubMed-9655938; Nilden F., Baeckstroem A., Bark C.; Molecular Cloning and characterisation of a mouse gene encoding an isoform of the neuronal cyclin-dependent kinase 5 (CDK5) activator."; Biochim. Biophys. Acta 1398:371-376(1998).
                                                                                                                                                                                                                                                                                                                                                                                                          "Comparative analysis of gene expression of the cyclin-dependent
Kinase 5 (cdk5) activators p35 and p39 in the rat CNS and embryonic
development revealed distinctive overlap with cdk5.";
Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
                                                                              CD5S_MOUSE STANDARD; PRT; 369 AA.
CD5S2_MOUSE STANDARD;
15-JUL-1999 (Rel. 38, Last sequence update)
15-JUL-1999 (Rel. 38, Last annotation update)
Cyclin-dependent kinase 5 activator 2 precursor (CDK5 activator 2)
CYCLIn-dependent kinase 5 regulatory subunit 2) (P39) (P39).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -!- SIMILARITY: BELONGS TO THE CYCLIN-DEPENDENT KINASE 5 ACTIVATOR
                                                                                                                                                                                                             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CYCLIN-DEPENDENT KINASE 5 ACTIVATOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -!- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A REGULATORY
-> G (IN REF. 2).
DA4EB1FD4D48CE66 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -> V (IN REF.
-> G (IN REF.
-> G (IN REF.
-> W (IN REF.
-> W (IN REF.
-> G (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POLY-LYS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-SER
                                                                                                                                                                                                                                                                                                                                                                                                Zheng M., Leung C.L., Liem R.K.H.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            38914 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; U90267; AAC53595.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; AF016393; AAB69709.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MGD; MGI:1330828; Cdk5r2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              163
351
91
172
191
197
220
                                                                                                                                                                                                 Mus musculus (Mouse).
                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            369 AA;
                                                                                                                                                                                                                                        NCBI_TaxID=10090;
                                                                                                                                                                                 CDK5R2 OR NCK5AI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             75
1155
348
348
91
1172
1191
120
220
220
253
                                                                                                                                                                                                                                                                                                                                                                                  STRAIN=129/SVJ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SUBUNIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROPEP
                                                                  CD5S_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CHAIN
                                                    RESULT
```

Length 369;

DB 1;

Score 80; DB Pred. No. 6.2;

5.9%;

Best Local Similarity

Query Match

43 DKIGDHHTHEHDESVSHVGLQAHFETW-LQMHHATKQEV--VRYQAYLQSRLGNYLPPMS 99

5.9%; Score 80.5; DB 1; Length 474; 23.2%; Pred. No. 7.6; tive 28; Mismatches 56; Indels 49

Best Local Similarity 23.28 Matches 39; Conservative

ŏ

Query Match

٠<u>.</u>

45; Gaps

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
   œ
                                                                                                                                         250 DELASAAELQAAFLTCL-----YLAY--SYMGNEISYPLKPFLVEPDKERFWQ 295
                                                                                                                                                                                                                                                         54 DESVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLP-PMSQLLT---TARSWQ 109
                                                                                                                                                                                                                110 ACGHEPYQLPPE-----HLWGQIVPTLHLYQDLKSRGIL-----PANTQIRSVYR 154
42; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Oxidoreductase; Chloroplast; Transit peptide; Nitrate assimilation;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -i- PATHWAY: SECOND STEP IN NITRATE ASSIMILATION (DENITRIFICATION).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    01-NOV-1988 (Rel. 09, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Ferredoxin--nitrite reductase, chloroplast precursor (EC 1.7.7.1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Streptophyta; Embryophyta; Tracheophyta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              INDUCTION: BY NITRATE.
SIMILARIY: THE C-TEBENIAL DOMAIN IS A 4FE-4S/SIROHEME DOMAIN
FOUND IN NITRITE REDUCTASES (EC 1.6.6.4 AND EC 1.7.7.1) AND
SULFITE REDUCTASES (EC 1.8.1.2 AND EC 1.8.7.1).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         confers nitrate inducibility on GUS gene expression in transgenic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             E., Dunne W., Schneiderbauer A., de Framond A., Rastogi R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Back E., Burkhart W., Moyer M., Privalle L., Rothstein S.;
"Isolation of cDNA clones coding for spinach nitrite reductase:
complete sequence and nitrate induction.";
Mol. Gen. Genet. 212:20-26(1988).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "Isolation of the spinach nitrite reductase gene promoter which
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -!- CATALYTIC ACTIVITY: Anmonia + H(2)0 + OH(-) + 3 oxidized ferredoxin = nitrite + 3 reduced ferredoxin.
-!- COFACTOR: THIS ENLYME CONTAINS ONE SIROHEME AND ONE 4FE-45 IRON-SULEUR CENTER AS PROSTHETIC GROUPS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota, Viridiplantae, Streptophyta; Embryophyta; Tracheo
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
   14; Mismatches 44; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STRAIN-CV. DARK GREEN BLOOMSDALE; TISSUE-Leaf;
MEDLINE-91329742; PubMed-1868226;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           594 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE-88232431; PubMed-3163766;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Biol. 17:9-18(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-NOV-1988 (Rel. 09, Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR000660; Nir_Sir.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pfam; PF01077; NIR_SIR; PRINTS; PR00397; SIROHAEM. PROSITE; PS00365; NIR_SIR; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X07568; CAA30453.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CAA34893.1; -.
                                                                                                                                                                                                                                                                                                                                                             155 NPELNQCAGGAAMSKHLT 172
                                                                                                                                                                                                                                                                                                                                                                                                                             350 SSARDSCATGA---KHWT 364
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Spinacia oleracea (Spinach)
       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; X17031; CAA341
PIR; S16603; S16603
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCBI_TaxID=3562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rothstein S.J
       38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NIR_SPIOL
P05314;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Plant Mol.
           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NIR_SPIOL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    A STANDAR STAN
                                                                                                                                                 g
                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                 δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                a
                                                                                                                                                                                                                       ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
                                                                                                                                                                                                                               13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no
                                                                                                                                                                                                                                                                                                                                                                                                   ---GAAMS 168
                                                                                                                                                                                                                                                                                                                                                                                                                        169 KHLT-NSAIDIWVPDLE-----IKSQALYELQNRLCQYWLEHGENQNFGLGLYATG-- 218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    238 QFVTANSRGNLSITNLPRKWNPCVIGSHDLYEHPHINDLAYMPATKNGKFGFNLLVGGFF 297
                                                                                                                                                                                                                                                                                                                                    77 KQEVVRYQ-------AYLQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPPE 121
                                                                                                                                                                                                                                                                                                77 NPAEKVKIEKDPMKLFIEDGISDLATLSMEEVDKSKHNKDDIDVRLK-----WLGLFHRR 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Oliver K.) Harris D., Parkhill J., Barrell B.G., Rajandream M.A.; Submitted (JAN-1998) to the EMBL/GenBank/DDBJ databases.
-!- FONCTION: THE REDD PROTEIN IS PROBABLY ONE OF SEVERAL DELICATELY BALANCED REGULATORY FACTORS THAT CONTRIBUTE TO THE CONTROL OF THE BIOSYNTHESIS OF THE ANTIBIOTIC UNDECYLFRODIGISIN (RED) IN
                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                             27 NP--QVSPIKTPSVLITKDKIGDHHT-----HEHDESVSHVGLQAHFETWLQMHHAT 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Narva K.E., Feitelson J.S.;
"Nucleotide sequence and transcriptional analysis of the redD locus
of Streptomyces coelicolor A3(2).";
J. Bacteriol. 172:326-333(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -!- SIMILARITY: BELONGS TO THE AFSR/DNRI/REDD FAMILY OF REGULATORS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
                             FERREDOXIN--NITRITE REDUCTASE.
IRON-SULEUR OR HEME IRON (POTEWTIAL).
IRON-SULEUR OR HEME IRON (POTEWTIAL).
IRON-SULEUR OR HEME IRON (POTEWTIAL).
                                                                                                                                                                                                                                                                                                                                                                 132 KHHYGRFMMRLKLPNGVTTSEQTRYLASVIKKYGKDGCADVTTRONWO------
                                                                                                                              i -> v (PROBABLE ALLELIC VARIATION).
                                                                                                                                                                                                                                 91;
                                                                                               (4FE-4S) AND SIROHEME
                                                                                                                                                                                               DB 1; Length 594;
                                                                                                                                                                                                                                 97; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     219 -----AIHLDT------20GFRKWGAQFSETNSICR 242
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                298 SIKRCEEAIPLDAWVSAEDVVPVCKAMLEAFRDLG--FRGNRQKCR 341
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Streptomyces coelicolor.
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
                                                                                                                                                  2B3DCAAC16DE06A3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                     122 HLWGQIVPTL-HLYQDLKSRGILPANTQIRSVYRNPELNQCAG-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -AUG-1990 (Rel. 15, Created)
-AUG-1990 (Rel. 15, Last sequence update)
5-DEC-1998 (Rel. 37, Last annotation update)
                                                                                                              (BY SIMILARITY
                                                                                                                                                                                                                                   35; Mismatches
                                                                                                                                                                                                                 19;
                                                                                                IRON-SULFUR
                                                                                                                                                                                               5.7%; Score 77.5; 22.0%; Pred. No. 19
                 CHLOROPLAST
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=A3(2);
MEDLINE=90094237; PubMed=2294088;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Transcriptional regulator redD.
                                                                                                                                                  66394 MW;
                                                                                                                                                                                                                                   63; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STANDARD;
               32
594
473
479
514
                                                                                                                                  221
                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
                                                                                                                                                  594 AA;
Heme; Iron-sulfur;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  .COELICOLOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REDD OR SC2E9.18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=1902;
                               33
473
479
514
518
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=A3(2);
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REDD_STRCO
P16922;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15-DEC-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-AUG-1990
                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                     Query Match
Best Local
                 TRANSIT
                                                                                                                                    VARIANT
                                                                  METAL
METAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 10
                                   CHAIN
                                                                                                    METAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          REDD_STRCO
                                                   METAL
                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                       a
                                                                                                                                                                                                                                                                                                                                                                     QQ
                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
   KW
FT
FT
FT
SO
SO
                                                                                                                                                                                                                                                                                                                                         á
```

```
ALA2_ARATH
ID ALA2_ARATH
AC P98205;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   clones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DOMAIN
                                                                                                      165
                                                                                                                                                                                RESULT
                                           ò
                                                                      g
                                                                                                      δ
modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
                                                                                                                                                                                                              70 LOMHHATKQEVVRYQA-----YLQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPPEH 122
                                                                                                                                                                                                                                   123 LWGQIVPTLHLYQDLKSRG--ILPAN----TQIRSVYRNPEL-----NQCAGGA 165
                                                                                                                                                                                                                                                                                            54; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                   166 AMSKHLTNSAIDIWVPDLEIK---SQALYELQNRLCQYWLEHGENQNFGLGLXATG 218
                                                                                                                                                                                                                                                                                                                                                        70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEDLINE-87282264; PubMed-3039043; Patel K.R., Smith K.T., Campo M.S.; "The nucleotide sequence and genome organization of bovine
                                                                                                                                                         DB 1; Length 350;
                                                                                                                                                                                        49; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 5.7%; Score 77; DB 1; Length 506; Best Local Similarity 21.1%; Pred. No. 17; Matches 38; Conservative 20; Mismatches 52; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Bovine papillomavirus type 4.
Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
                                                                         EMBL; AL021530; CAA16486.1; -
Antibiotic biosynthesis; Transcription regulation.
SEQUENCE 350 AA; 37796 MW; 573C96C2203EFBBB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                57902 MW; F1C5DFDB54FA681E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (Rel. 08, Created)
(Rel. 08, Last sequence update)
(Rel. 34, Last annotation update)
                                                                                                                                                                                    32; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                 506 AA
                                                                                                                                                         Score 77; DB Pred. No. 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL; X05817; -; NOT_ANNOTATED_CDS. EMBL; D00146; BAA00101.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    papillomavirus type 4.";
J. Gen. Virol. 68:2117-2128(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 InterPro; IPR002210; PV_capsid_L1.
Pfam; PF00500; late_protein_L1; 1.
PRINTS; PR00865; HPVCAPSIDL1.
ProDom; PD000544; PV_capsid_L1; 1.
                                                                                                                                                      Match 5.7%;
Local Similarity 23.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Coat protein; Late protein.
SEQUENCE 506 AA; 57902 M
                                                              EMBL; M29790; AAA88556.1;
                                                                                                                                                                                   41; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                              STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Major capsid protein L1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PIR; B26214; PIWLB4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=10562;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Papillomavirus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-AUG-1988 (
01-AUG-1988 (
01-OCT-1996 (
                                                                                                                                                                                                                                                                                                                                                                                                                                           VL1_BPV4
P08341:
                                                                                                                                                      Query Match
                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 1
VL1_BPV4
   ò
                                                                                                                                                                                                                                           q
                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                    à
```

---QLPPEHLW--- 124

79 EVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGHEPY-

<del>--</del>

ò

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the Buropean Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Structural analysis of Arabidopsis thaliana chromosome 5, I. Sequence features of the 1.6 Mb regions covered by twenty physically assigned
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               125 ---GQIVPT----LHLYQDLKSRGILPANTQIRS------VYRNPELNQCAGG-- 164
                                                        261 GRVGDDIPTGESGSPYFLPATGRGPLPSSVYIGSPSGSLVSSDQQIYNRPFWIQRAQGSN 320
                                                                                          -- AAMSKHLTNSAIDIWVPDLEIKSQALYELQNRLCQYWLEHGENQNFGL 212
                                                                                                                  321 NGMCWNNELFVTAVDSTRGTNFSISVHTTDPEVKPQETYTATK--FKHYLRHVEEWDLSL 378
                                                                                                                                                                                                                                                                                                             Arabidopsis thaliana (Mouse-ear cress),
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PROSTIE: PS01154; ATTACKEE. ELE2; 1.
Hydrolase; Prosmembrane; Phosphorylation; ATP-binding; Metal-binding;
Magnesium; Multigene family.
                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN=CV. COLUMBIA;
MEDLINE-97471969; PubMed=9330910;
Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
Miyajima N., Tabata S.;
                                                                                                                                                                                                                    16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Potential phospholipid-transporting ArPase 2 (EC 3.6.3.1)
ALA2 OR AT5G44240 OR MLN1.17.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EXTRACELLULAR (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     EXTRACELLULAR (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                                              PRT; 1107 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AB005239; BAB10991.1; ALT_SEQ.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   InterPro; IPR001757; E1-E2_ATPase.
InterPro; IPR001454; Hydrolase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CATATPASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PF00702; Hydrolase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA Res. 4:215-230(1997).
                                                                                                                                                                                                STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONCEPTUAL TRANSLATION.
                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=3702;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRINTS; PR00119;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       34
334
334
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
TRANSMEM
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FRANSMEM
```

Wed Jul 31 08:20:00 2002

```
97 KANEK-----EVWI----VKQGIKKHIQAQDIQVGNIVWLRENDEVPCDLVLLGTSDP 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            168 SKHLTNSAIDIWVPDLEIK----SQALYE--LQNRLCQYWLEHGENQN-----FGL 212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -------DTQGFR 228
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   238 SVYTGNOTKLGMSRGIAEPKLTAMDAMIDKLTGAIFVFQIVVVLVLGIAGNVWKDTEARK 297
                                                                                                                                                                                                                                                                                                                                                                                                     56 SVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNY------LPPMSQLLTTARS 107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                108 WQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAM 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             146 QGVCYVETAALDGE------TDLKTR-VIP-----SACVGIDLE 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         halodurans and genomic sequence comparison with Bacillus subtilis."; Nucleic Acids Res. 28:4317-4331(2000).
                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                         5 NQYFITTLISSMLVAC----SAPIPTNPQVSPIKTPSVLI-----TKDKIGDHHTHEHDE 55
                                                                                                                                                                                                                                                                                                                                                                     46 NOYF-----LLIACLQLWSLITPVNP--ASTWGPLIFIFAVSASKEAWDDYHRYLSDK 96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Maeno G., Sasaki R., Masui N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           "Complete genome sequence of the alkaliphilic bacterium Bacillus
                                                                                                                                                                                                                                                                                                             85; Indels 124;
                                                                                                                                                                                 PHOSPHORYLATION (BY SIMILARITY).
MAGNESIUM (BY SIMILARITY).
MAGNESIUM (BY SIMILARITY).
MA; 31F7722E9653C96F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
16-0CT-2001 (Rel. 40, Last annotation update)
Pantothenate Kinase (EC 2.7.1.33) (Pantothenic acid Kinase).
                                                                                                                                                                                                                                                                           Query Match 5.7%; Score 77; DB 1; Length 1107; Best Local Similarity 19.2%; Pred. No. 45;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STRAIN-C-125 / JCM 9153;
MEDLINE-21051282; PubMed-11058132;
Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., M.
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             phosphopantothenate.
-!- PATHWAY: Coenzyme A (COA) biosynthesis; first step.
POTENTIAL.
EXTRACELLULAR (POTENTIAL).
                                                                         EXTRACELLULAR (POTENTIAL).
                                                                                                                                       EXTRACELLULAR (POTENTIAL).
                           POTENTIAL. CYTOPLASMIC (POTENTIAL).
                                                                                                        CYTOPLASMIC (POTENTIAL).
                                                                                                                                                                      CYTOPLASMIC (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                                                                                                                                                             40; Mismatches
                                                                                                                          POTENTIAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                  124835
                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacillus halodurans.
                                                                                                                                                                                                                                     AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        229 KWGAQFSE 236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      298 QWYVQYPE 305
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NCBI_TaxID=86665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            213 GLYA-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COAA OR BH2875
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Horikoshi K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COAA_BACHD
                                                                                                                                                                                                                                                                                                                 26.
                                            DOMAIN
TRANSMEM
DOMAIN
                DOMAIN
TRANSMEM
                                                                                                                                                                                                                                    SEQUENCE
                                                                                                                          TRANSMEM
                                                                                                                                                         TRANSMEM
                                                                                                                                                                       DOMAIN
MOD_RES
METAL
                                                                                           FRANSMEM
   TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9K8X7;
                                                                                                             DOMAIN
                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 13
COAA_BACHD
                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
                                                                                                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δý
```

```
this SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation-the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                     This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its was by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDLINE=96357133; PubMed=8760884;
Inoue A., Takahashi K., Kimura M., Watanabe T., Morisawa S.;
Inoue A., Takahashi K., Kimura M., Watanabe T., Morisawa S.;
"Molecular cloning of a RNA binding protein, S1-1.";
Nucleic Acids Res. 24.2990-2991;
I- SUNCTION: NOT KNOWN. BINDS TO RNA HOMOPOLYMERS, WITH A PREFERENCE
FOR POLY(G) AND POLY(U) AND LITTLE FOR POLY(A).

-!- SUBCELLUIAR LOCATION: Nuclear.
-!- SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
-!- SIMILARITY: CONTAINS 2 RNA RECOGNITION MOTIFS (RRM).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           109 QACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMS 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     169 KHLT-NSAIDIWVP---DLEIKSQALYELONRLCQYWLEHGENON-FGLGLYATGAIHLD 223
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RNA-binding protein 10 (RNA binding motif protein 10) (S1-1 protein).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBL_TaxID=10116;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             5 EACDFFPYTVLSRSQW-----KSLRKASSLPINEQ-----ELEQLVG---LN 43
                                                                                                                                                                                                                                                                                                                                                                                                                                                27; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                          5.6%; Score 76.5; DB 1; Length 316; 26.4%; Pred. No. 11;
                                                                                                                                                                                                                                                                                    Transferase, Kinase, ATP-binding, Coenzyme A biosynthesis; Complete proteome.
SUBCELLULAR LOCATION: Cytoplasmic (Probable). SIMILARITY: BELONGS TO THE PANTOTHENATE KINASE FAMILY.
                                                                                                                                                                                                                                                                                                                             ATP (POTENTIAL).
3196BA013E3B0BB3 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-0CT-2001 (Rel. 40, Created)
16-0CT-2001 (Rel. 40, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               852 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                           26.4%; Pred. no. 11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                           EMBL; AP001516; BAB06594.1; -.
                                                                                                                                                                                                                                                                                                                                                  36416 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        EMBL; D83948; BAA12144.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                    32; Conservative
                                                                                                                                                                                                                                              InterPro, IPR001324; PRK. Pfam; PF00485; PRK; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Rattus norvegicus (Rat).
                                                                                                                                                                                                                                                                                                                               102
                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                  316 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               103 T 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          224 T 224
                                                                                                                                                                                                                                                                                                                               NP_BIND
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RBMA RAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RBM10
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ô
```

```
Search completed: July 30, 2002, 16:18:37 Job time: 204 sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    163 GGAAMSKHLTNSAID----
                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                  Matches
            οy
                                                                                                                                                                                                                                                                                                                                                                                                                                    q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                             ر.
د
                                                                                                                                                                                                                                                                                                                                                                                                                                         450 TIMSPAVLKSELOSPTHPSSALPPATSPTAPESYSQYPVPDVSTYQYDETSGY----- 502
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   67 ETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPPEHLWGQ 126
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       127 IVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDIWVPDLEIK 186
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      halodurans and genomic sequence comparison with Bacillus subtilis.";
Nucleic Acids Res. 28:4317-4331(2000).
-!- FUNCTION: This protein is involved in the repair of mismatches in DNA. It is possible that it carries out the mismatch recognition
                                                                                                                                                                                                                                                                                                                                                                                            53; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                 11 TLISSMLVACSAPIPTNPQ--VSPIKTPSV--LITKDKIGDHHTHEHDESVSHVGLQAHF 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ----YYDPQTGLYYDPNSQYYYNAQSQQ-----YYDPQTGLYYDPNSQYYYNAQSQQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
Horikoshi K.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             "Complete genome sequence of the alkaliphilic bacterium Bacillus
                                                                                                                                                                                                                                                                                                                                                               / Match 5.6%; Score 76.5; DB 1; Length 852; Local Similarity 19.6%; Pred. No. 36; nes 36; Conservative 30; Mismatches 65; Indels 5
                                                                                                                                                                                                                                                                                                                                                                                           65; Indels
                                                                                                                                                                                                                                                                                                                         D6D2EECFEECBE189 CRC64;
                                                                                                                                                                                                                                           protein; zinc-finger; Repeat.
                                                                                                                                                                                                                                                                                  RNA-BINDING (RRM) 2.
C2H2-TYPE (ATYPICAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus/Staphylococcus group; Bacillus.
                                                                                                                                                                          PROSITE; PS50102; RRM; 2.,
PROSITE; PS00030; RRM_RNP_1; 2.
PROSITE; PS01358; ZF_RANBP2_1; 1.
PROSITE; PS50199; ZF_RANBP2_2; 1.
PROSITE; PS50157; ZINC_FINGER_C2H2_2; UNKNOWN_1.
RNA-binding; Nuclear protein; Zinc-finger; Repea
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              01-WAR-2002 (Rel. 41, Created)
01-WAR-2002 (Rel. 41, Last sequence update)
01-WAR-2002 (Rel. 41, Last annotation update)
DNA mismatch repair protein muts.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         865 AA.
                                                                                                                                                                                                                                                                                                                                                                                         30; Mismatches
                                                                                                                                                                                                                                                                       RANBP2-TYPE.
                                                                                                                                                                                                                                                                                                              G-PATCH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN=C-125 / JCM 9153;
MEDLINE=20512582; PubMed=11058132;
InterPro; IPR000467; G_patch.
InterPro; IPR000504; RRM.
InterPro; IPR000822; Znf-C2H2.
InterPro; IPR001876; Znf-RanBP.
Pfam; PF001585; G-patch; 1.
Pfam; PF00076; zrm; 2.
Pfam; PF00064; zf-C3H2; 1.
Pfam; PR00064; zf-C3H2; 1.
SMART; SM00443; G-patch; 1.
SMART; SM00455; Znf-C2H2; 1.
SMART; SM00355; ZnF_C2H2; 1.
PROSITE; PS50174; G_PATCH; 1.
PROSITE; PS50174; G_PATCH; 1.
                                                                                                                                                                                                                                                                                                                         94387 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
SEQUENCE 125 / JCM 9153;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Bacillus halodurans.
                                                                                                                                                                                                                                                                                                                        852 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               503 -----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MUTS OR BH2369.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           187 SQAL 190
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      581 ARSL 584
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MUTS_BACHD
                                                                                                                                                                                                                                                                                                                                                                Query Match
                                                                                                                                                                                                                                                                                                                         SEQUENCE
                                                                                                                                                                                                                                                                                               ZN_FING
DOMAIN
                                                                                                                                                                                                                                                                    ZN_FING
                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 15
MUTS_BACHD
                                                                                                                                                                                                                                                                                                                                                                                        Matches
    QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            οq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          οy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              114 EPYQLPPEHLWGQI----VPTLHLYQDLKSRGILPANTQIRSVYRNPELNQ-----CA 162
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           54 DESVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGH 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           282 DETVTAMGGRL-LKQWVERPLLSKKEIERRQGLVQSFLDHYF-EREELRDELR----- 332
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       --IWVPDLEIKSOALY 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         391 SLVDLLEH---SLVDDPPVSIKEGGMIRDGFHKELDTYRDASRNGKSWIAELEQKEREAT 447
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               192 ELQNRLCQYWLEHGENQNFGLGLYATGA-IHLDTQGFRKWGAQFSETNSICRHVLPKNK 249
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            step. This protein has a weak ArPase activity (By similarity).
-!- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               :99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             'Match 5.6%; Score 76; DB 1; Length 865; Local Similarity 22.6%; Pred. No. 41; less 54; Conservative 40; Mismatches 79; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROSITE: PS0046; DNA_MISNATCH_REPAIR_2; 1.
DNA repair; ATP-binding; DNA-binding; Complete proteome.
NP_BIND 605 612 ATP (POTENTIAL).
SEQUENCE 865 AA; 98367 MW; A00D4346163DE51C CRC64;
                                                                                                                                                                                                                                                                                                                                                                                     EMBL; AP001515; BAB06088.1; -.
InterPro; IPR000432; MutS_C.
InterPro; IPR0002863; MutS_N.
InterPro; IPR002019; Urease_beta.
Pfam; PF01624; MutS_C; 1.
Pfam; PF01624; MutS_N; 1.
ProDom; PD001263; MutS_C; 1.
SMART; SM00534; MUTSaC; 1.
```

^

```
July 30, 2002, 16:15:33 ; Search time 26.57 Seconds (without alignments) 1627.728 Million cell updates/sec
                                                                                                                                                                                                                         1355
1 MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250
                                                                                                                                                                                                                                                                                                                                                                                                562222
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
                                                                                                                                                                                                                                                                                                                                                                                      Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                 562222 seqs, 172994929 residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                               OM protein - protein search, using sw model
                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                                                                                                                                                                                                                                                                                                                                                                                            Minimum DB seq length: 0
Maximum DB seq length: 200000000
                                                                                                                                                                                                       US-09-674-779-2
                                                                                                                                                                                                                       Perfect score:
                                                                                                                                                                                                                                                                                     Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Database :
                                                                                                                                                                                                                                              Sequence:
                                                                                                                                                                                                                                                                                                                                                 Searched:
                                                                                                                      Run on:
```

| SpTREMBL\_19:\*
| Sp\_archea:\*
| Sp\_archea:\*
| Sp\_bacteria:\*
| Sp\_bacteria:\*
| Sp\_human:\*
| Sp\_human:\*
| Sp\_nowertebrate:\*
| Sp\_nowertebrate:\*
| Sp\_nowertebrate:\*
| Sp\_nowertebrate:\*
| Sp\_nowertebrate:\*
| Sp\_phage:\*
| Sp\_phage:\*
| Sp\_phage:\*
| Sp\_phage:\*
| Sp\_rotrebrate:\*
| Sp\_vortebrate:\*
| Sp\_vortebrate:\*
| Sp\_vortebrate:\*
| Sp\_rotrus:\*
| Sp\_rotrus:\*
| Sp\_archeap:\*
| Sp\_archeap

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result Query

No. Score Match Length DB ID

1 375 27.7 240 2 Q93D15
2 122 9.0 562 16 Q98B14 Q98B17 Lizoblum ID

4 102.5 7.9 622 16 Q98B18 Q98B18 Q98B18 Lizoblum ID

5 101 7.5 216 G98B18 Q98B2 Caulobacter G98.5 7.3 193 16 Q98B18 Q98B18 Q98B2 Caulobacter ID

8 89.5 6.6 1377 5 Q94NZP3 Q99V28 Garosophila Q98.5 6.6 1377 5 Q94NZP3 G99V28 Grosophila Q98B18 Gos G1377 5 Q94NZP3 G99V28 Grosophila G98B18 Gos G1377 5 Q94V2P3 G99V2B1 G99V2B18 G

032431 acinetobact 043981 acinetobact 051823 thiocapsa r 099w22 drosophila 09ck19 pasteanralla 09cw15 drosophila 092w16 passiflora 09nh67 drosophila 0991p4 rhizobium 1 09vjh6 drosophila 09jvl4 neisseria m 09jvl6 drosophila 09tvl6 drosophila 09tvl7 drosophila 09svg drosophila 09svg drosophila 09f21 homo sapien 09f21 homo sapien 09f21 homo sapien 09f21 homo sapien 09gd1 mytella fas 09gwc drosophila 09gwc arabidopsis 093wc arabidopsis	ALIGNMENTS  PRT; 240 AA.  Created) Last sequence update) Last annotation update) Last annotation watere) Last annotation watere) Last annotation watere) Last annotation watere) Last sequence update) Last sequence update update) Last sequence update upda	tch 27.7%; Score 375; DB 2; Length 240; 81 Similarity 35.9%; Pred. No. 2.2e-28; 82 Conservative 45; Mismatches 80; Indels 34; Gaps 10; TLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSH 59  11
032431 043981 0551823 095012 096011 098M5 099M67 099M67 099M6 09123 09123 091589 091689	ALIGNMENT  PRT; 240  19, Created) 19, Last sequenc 19, Last annotat NOTEIN.  2; gamma subdivis 2; gamma Subdivis 2; gamma Subdivis 3; gamm	Score 3. Pred. No. 5; Missmann Str. KTPSVLITI VQAYLQSRI I::  I::  YEQELSRI LPANTQIR
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	L1 19, 11, 19, 11, 19, 11, 19, 11, 19, 11, 19, 11, 11	7%; 4; 4; 4; VSPII) EVVR: 5; SRGIII: 1:
5111 1218 3318 3318 6532 4430 1062 1062 1062 4448 1130 1150 1130 1130 1130 1130 1130 1130	PRELIMINARY;  (Trembrel. 19, Ci (Trembrel. 19, Li (Trembrel. 19, L	tch al Similarity 35.9%; Pr 89; Conservative 45; TLISSMLVACSAPIPTNPQVSPIKTP SIPIMMVGCTTAPQ VGLQAHFETWLQMHHATKQEVVRYQA
	PRELI (Trew (Trew (Trem L 27.4 L 27.4 L 27.4 L 27.4 L 27.4 L 27.4 SS 2977, SS 2977, SS 2987, SS 998, SS 998, S	Similarity 9; Conserv SSMLVACSAPI  :   :  :  -  -  -  -  -  -  -  -  -  -  -  -  -
83.5 83.5 83.5 83.5 83.5 83.5 82.5 82.5 81.5 81.5 81.5 80.5 80.5 79.5	ALIG 0127 1 023015, PRELIMINARY; PRT; 023015, 01-DEC-2001 (TrEMBLrel. 19, Created 01-DEC-2001 (TrEMBLrel. 19, Last so 01-DEC-2001 (TrEMBLrel. 19, Last so NCHETICAL 27,4 KDA PROTEIN. Acinetobacter Sp. ADP1. Bacteria; Proteobacteria; gamma sub Acinetobacter. NCBI_TAXID=6297; [1] SEQUENCE FROM N.A. SEQUENCE FROM N.A. SEQUENCE FROM N.A. SEQUENCE FROM N.A. SEQUENCE ROM N.A. SEQUENCE ROM N.A. SEQUENCE ACINETORIAL SPECIAL SECONDALIA SECONDALIAL SECONDALIA SECONDALIA SECONDALIA SECONDALIA SECONDALIA SECONDALIA SECONDALIA SECONDALIA SECONDALIA SEQUENCE 240 AA; 27436 MW; EOFF.	Ma Loc es 11 12 60 57
, , , , , , , , , , , , , , , , , , ,	to m	Best Match 1
	RESERVED BY CONTRACT OF CONTRA	OY OY OY OY OY

a

```
Phyllobacteriaceae; Mesorhizobium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                  Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       098988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MLR6494
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q989B8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q989B8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ó
                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                           δλ
  ŏ
                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                              Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F., Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G., Barloy-Hubler F., Buhrmester J., Cadieu E., Capela D., Chain P., Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F., Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M., Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D., Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V., Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B., Ramsperger U., Surzycki R., Thebault P., Vandenbol M., The composite genome of the legume symbiont Sinorhizobium meliloti.", Science 293:668-672(2001).

EMBL, AL591791: CAC47333.1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               74 HATKQEVVRYQAYLQSRLGNY-LPPMSQLLTTARSWQACGHEPYQLPPEHL---WGQIVP 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        130 TLHLYQDLKSRGILPANTQIRSVYRNPELN----QCAGGAAMSKHLTNSAIDIWVPDLE 184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        180 V-PDLBIKSQALYELQN---RLCQYWLEHGENONFGLGLYATGAIHLDTQGFRKWGAQFS 235 : | :: | :: | :: | | | | | | | |
 115 NRELWNNAIPTLRVLKYLVAAKIL-TDFEVTSVYRDLPLNQCAGGANSSRHLFNSAIDFR 173
                                                        174 IGPEIP-QPQDYAFIENTKFKLCOFWNQHGQSLNMGLGLYSSGQIHIDTQGYRTWGPDHS 232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20 HTKEKAQITYK-----RNGRYDQKGLQQINRFLRDWRR--NEPTKMDPRLLDLVW----- 67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Rhizobium loti (Mesorhizobium loti).
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
                                                                                                                                                                                                                                                                                Rhizobium meliloti (Sinorhizobium meliloti),
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Sinorhizobium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        9.0%; Score 122; DB 16; Length 562; 25.4%; Pred. No. 0.0024; ive 29; Mismatches 55; Indels 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               185 IKSQALYELQNRLCQYWLEHGENQNFGLGLYATGA---IHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         122 LKT--LREVGMKF-----QVGGVGYYPTSGSPFVHMDVGGVRAW 158
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  12C41231BACD3422 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                               01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Hypothetical protein; Complete proteome. SEQUENCE 562 AA; 59535 MW; 12C41231B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                     MEDLINE=21368234; PubMed=11474104;
                                                                                                                                                                                                                                                         HYPOTHETICAL PROTEIN SMC04010. SMC04010.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (TrEMBLrel. 18, (TrEMBLrel. 18, (TrEMBLrel. 18,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 25.41
Matches 43; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-0CT-2001 (TrEMBLrel. 1
01-0CT-2001 (TrEMBLrel. 1
01-0CT-2001 (TrEMBLrel. 1
MLL2999 PROTEIN
MLL2999
                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                            SECUENCE FROM N.A.
                                                                                       236 ETNSICRH 243
                                                                                                                  233 RNSSMCNY 240
                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=382;
                                                                                                                                                                                                  Q92JX1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 098H74;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Q98H74
                                                                                                                                                                                       092JX1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
Q98H74
                                                                                                                                                           RESULT
                                                                                                                                                                          Q92JX1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AC DT DT DT OC OC OC
                                                                                                                                                                                       ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              pp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                           a
                                                                                        δλ
                                ò
```

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Pakeuchi Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabbta S., "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizoblum loti.":
                                                                    Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S., "Complete genome structure of the nitrogen-fixing symbiotic bacterium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     72 MHHATKQEVV--RYQAYLQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIVP 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                128 VPTLHLYQDLKSRGILPANTQIRSVYRNPELN----QCAGGAAMSKHLTNSAIDIWVPD 182
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    76 ----EAYRESGATDYI----QVVCGYRSPATNSMLRSRSRGVAEKSQHMLGKAMDFYIPG 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     72 MHHATKQEVVRYQAYLQSRLGNYLPP-MSQLLTTARSWQACGHEPYQLPPEHL---WGQI 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Rhizobium loti (Mesorhizobium loti).
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Phyllobacteriaceae; Mesorhizobium.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 16; Length 523;
                                                                                                                                                                                                                                                                                                                                                                                                                              DB 16; Length 622;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      183 LEIKSQALYELQNRLCQYWLEHGENQNFGLGLYATGA---IHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  523 AA; 55318 MW; FA6196E5665414C4 CRC64;
                                                                                                                                                                                                                                                                                                                                               622 AA; 65402 MW; 1E4C6B5F8C291655 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               61;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7.6%; Score 102.5; DE 25.4%; Pred. No. 0.17; iive 29; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                              7.9%; Score 106.5; DB 23.4%; Pred. No. 0.086;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 27; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-0CT-2001 (TrEMBLrel. 18, Created)
01-0CT-2001 (TrEMBLrel. 18, Last sequ
01-0CT-2001 (TrEMBLrel. 18, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN=MAFF303099;
MEDLINE=21082930; PubMed=11214968;
                                             MEDLINE=21082930; PubMed=11214968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DNA RES. 7:331-338(2000).

EMBL, AP003009; BAB52779.1; -.

COMPLEE proteome.
                                                                                                                                                                                                                                                                  DNA Res. 7:331-338(2000).
EMBL; AP003000; BAB49992.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                           23.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        43; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                                                                                                                                                                                                        Complete proteome.
SEQUENCE 622 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
                                                                                                                                                                                                                                              Mesorhizobium loti
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MLR6494 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NCBI_TaxID=381;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-OCT-2001
01-OCT-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          40;
```

alpha subdivision; Rhizobiaceae group;

```
Bacteria; Proteobacteria; alpha su
Phyllobacteriaceae; Mesorhizobium.
                                            SEQUENCE FROM N.A.
                      NCBI_TaxID=381;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                095TM6
OCC
OCX
OCX
OCX
RRX
RRX
RRY
DR RTA
DR RTA
DR RTA
SQ SQ SQ
                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ΩD
                                                                                                                                                                                                                                                                      RN SEQUENCE FROM N.A.

RY SEQUENCE FROM N.A.

RY MEDLINB-21173698; PubMed=11259647;

RA Mierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,

RA Nierman W.C., Neldelberg J.F., Alley M.R.K., Ohlta N., Maddock J.R.,

RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,

RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,

RA Nollonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;

RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;

RComplete genome sequence of Caulobacter crescentus.",

Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10;
130 TLHLYQDLKSRGILPANTQIRSVYRNPELN----QCAGGAAMSKHLTNSAIDIWVPDLE 184
                  78 QEVVRYQAYLQSRLGNYLP-PMSQLLTTARSWQACGHEPYQLPPEHLWGQIVPTLHLYQD 136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             137 LKSRGILPANTQIRSVYRNPELNQCAGG-----AAMSKHLTNSAIDIWVPDLEIKSQALY 191
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28 PQVSPIKTPSVLIT------KDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATK 77
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -----VHTGEK 81
                                                                                                                                                                                                                          Caulobacter crescentus.
Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        62;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match 7.5%; Score 101; DB 16; Length 216; Best Local Similarity 22.7%; Pred. No. 0.073; Matches 50; Conservative 29; Mismatches 79; Indels 6;
                                                185 IKSQALYELQNRLCQYWLEHGENQNFGLGLYATGA---IHLDTQGFRKW 230
                                                                     131 LKK--LRDIGLKM-----QGGGVGYYPTSGSPFIHMDVGNVRHW 167
                                                                                                                                                                                                                                                                                                                                                                                                                                                              11 protein; Complete proteome.
216 AA; 23597 MW; 65D0901102F7BBD4 CRC64;
                                                                                                                                                           01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             || :||: |: |: 39 PQATPVPPTIVAATVASIDPPALKPAVDPRWVHLHN------
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              192 ELQNRLCQYWLEHGENQNFGLGLYAT-GAIHLDTQGFRKW 230
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               186 ALD-----LSVG-----GVGYYPTSNFVHVDVGPVRKW 213
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       193 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            098AI7;
01-0CT-2001 (TrEMBLrel. 18, Created)
01-0CT-2001 (TrEMBLrel. 18, Last seq
01-0CT-2001 (TrEMBLrel. 18, Last ann
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           MLL5985.
Rhizobium loti (Mesorhizobium loti).
                                                                                                                                          PRT;
                                                                                                                                                                                                  HYPOTHETICAL PROTEIN CC1512
                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY:
                                                                                                                                                                                                                                                 Caulobacter.
NCBI_TaxID=69394;
                                                                                                                                                  Q9A852;
                                                                                                                                      09A852
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Q98AI7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                        RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
                                                                      g
                       g
                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Watanabe A., Idesawa K., Ishkawa A., Kawashima K., Kimura T., Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.; "Complete genome structure of the nitrogen-fixing symbiotic bacterium
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pteryyota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54 DESVSHVGL--QAHFETW-----LQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTT 104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      105 ARSWQACGHEPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGG 164
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 84 EEEGQALGFVCLLGNAEDRWGTLVDNLHVLPTAKGRGV-------GRHLIRVAAG 131
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              54 DESVSHVGLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGH 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        25 DATVADVDLIAQLHIESWRNAYAGILSASYLAGPIVLDRQTVWRERFNSPAPNL-KVIVA 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               165 AAMSKHLTNSAIDIWVPDLEIKSQALYE------LQNRLCQYWLEHG 205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       : ::: || | : || | : || 132 WS-AENYPGYGLHLMYYEVNAPARAFYERMGGQVVARLPQSNPDGRIHVELCXYWPDSG 189
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ctapleton M., Drokstein P., Hong L., Agbayani A., Carlson J., Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R., Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M., Celniker S.; Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.

EMBL; AYO58679; AAL13908.1; -- SPGOUENCE 539 AA: 59662 MW; 423555844D084EFA CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               41;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7.3%; Score 98.5; DB 16; Length 193; 21.2%; Pred. No. 0.11; Live 27; Mismatches 73; Indels 41
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         73; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match 6.6%; Score 90; DB 5; Length 539; Best Local Similarity 27.2%; Pred. No. 2.8; Matches 41; Conservative 19; Mismatches 63; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                               193 AA; 21306 MW; DA78C07070694B5E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              539 AA
                                                                                                                                                                                                                                                                                                                                InterPro; IPR000182; Acetyltransf_GCN5
Pfam; PF00583; Acetyltransf; 1.
                      MEDLINE=21082930; PubMed=11214968;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT;
                                                                                                                                                                                                                                                                                                   EMBL; AP003008; BAB52343.1; -.
                                                                                                                                                                                                                                           Mesorhizobium loti.";
DNA Res. 7:331-338(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A. STRAIN=Y, CN BW SP;
STRAIN-MAFF303099;
                                                                                                                                                                                                                                                                                                                                                                                                  Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LD38816P.
                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q95TM6
Q95TM6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CG8793
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
```

9

. 9

PRT; 1377 AA.

PRELIMINARY;

```
09W2F3
  10;
                                                                                                                                                                                                                                                                                                                            "PTP-ER, a novel tyrosine phosphatase, functions downstream of Rasl to downregulate MAP kinase during Drosophila eye development.";
                                                                                                                                                                                                        Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  401 TIDWLMYHRKQNPYQVQPTHCSSTTQSSLDSDASLTPSLGDFELKSACSVDGGSKFGIGA 460
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   461 SLAPRSAHKHNOLLHSSSTNLKTLPECLTLVEFSSSGGGPKESPFKOKSMDLPMPTLQAK 520
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             292 PIKNSSQPSPLPKPKTPTIKSTKEKARSLDSAANESELSIV------VHNITESH 340
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----EVVRYQAYLQ------SRLGNYLPPMSQLLTTARSWQACG 112
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                113 ------HEPYQLPPEHLW-------GQIVPTLHLYQ------DLKSRGILPA 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ----LNQC------AGGAAMSKHLTNSAIDIWVPDLEIK 186
             114 EPYQLPPEHLWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAG------GA 165
                           :| | |:|| :: || ::| |: || 290 ---pvpnrllageivp----Vhytlrnmgiap----Ieeiylgcdnprcvslldqhsqmpl 339
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     23 PIPTNPQVSPI---KTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATKQ- 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        6.6%; Score 89.5; DB 5; Length 1377;
18.8%; Pred. No. 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                87; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1377 AA; 152974 MW; 9BBDDE009FF95074 CRC64;
                                                                                                                                                                 01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
PROTEIN TYROSINE PHOSPHATASE ERK (EC 3.1.3.48).
                                                                                                                                           PRT; 1377 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                42; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                              SMART; SM00194; PTPC; 1.
SMOSITE; PS00383; TYR_PHOSPHATASE_1; 1.
PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   521 TTVSTSSMNLL----QRRGSNHSLTLNLHSS 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             187 SQALYELQNRLCQYWLEHGENQNFGLGLYAT 217
                                                                           340 AMMSSLRNLSNDKLVKDKEIRGQRVYRLLNR 370
                                                             166 AMSKHLTNSAIDIWVPDLEIKSQALYELQNR 196
                                                                                                                                                                                                                                                                                                                                                                                   Flybase; FBG00016641; PTP-ER.
InterPro; IPR000387; TYR_phosphatase.
InterPro; IPR000242; TYT_prof_phphtase.
Pfam: PF00102; Y_phosphatase; I.
PRINTS; PR00700; PRTYPHPHTASE.
                                                                                                                                                                                                                                                                                                       MEDLINE=9932266; PubMed=10394362;
Karim F.D., Rubin G.M.;
                                                                                                                                                                                                                                                                                                                                                   MOI. Cell 3:741-750(1999).
EMBL; AF146594; AAD38688.1; -.
HSSP; Q06124; 2SHP.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local Similarity 18.8%
Matches 51; Conservative
                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                146 NTOIRSVYRNPE----
                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . . . . . 62
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Hydrolase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                        8X0Y60
                                                                                                                                            09Y0X8
                                                                                                                      8
                                                                                                                       RESULT
                                                                                                                                  09Y0X8
                                                                                                                                                        Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δŏ
                  ò
                                                              δŏ
```

RESULT Q9W2F3

```
RC STRINBERELELY.

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayue J.D.,

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayue J.D.,

RA Adams M.D., Celniker S.E., Holt R.A., Hoskins R.A., Galle R.F.,

George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Gorge R.A., Lewis S.E., Holt R.A., Champe M., Prelifer B.D.,

Button G.G., Wortuna J.R., Yandell M.D., Zhango Q., Chen L.X.,

RA Man K.H., Doyle C. Baxter E.G., Helf G., Nelson C.R., Miklos G.L.G.,

RA Adril J.F., Agbayani A., An H.-J., Andrews Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Ballew R.M., Basu A., Bandall M.D., Brokstein P., Brottler P.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottler P.,

RA C., Busan D.A., Buller H., Cadieu E., Center A., Chandra I.,

RA Cayler C., Cawley S., Downes M., Dugan Rochas S., Pleischmann W.,

RA Cherry J.M., Cawley S., Dobhike C., Davenport L.B., Davies P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Pleischmann W.,

RA Glodek A., Gong F., Gorell J.H., Gu Z., Galbar W. M., Glasser K.,

RA Harris N.L., Harrey D., Heiman T.J., Wei M. H., Libayam C.,

RA Harris N.L., Harrey D., Heiman T.J., Wei M. H., Libayam C.,

RA Jalali M., Kallah F., Karpen G.H., Ke Z., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Martei B. W. Milshina N.V. Mobarty C., Morris J., Puris N.,

RA Martei B. W., Murphy B., Murphy L., Muzny D.M., Nelson D.,

RA Martei B. W., Wurphy B., Murphy L., Weinsenbach J.,

RA Beinert K., Remigoton K., Sauders R., Sun B.,

Spica E., Speahiding A.C., Stapleton M., Strong R., Sun B.,

RA Harris A., Woodage T., Wolley K.C., Wu D., Yang S., Yao Q., Zheng L.,

RA Mang Z.-Y., Wassarman D.A., Weinstey M., Strong R., Shong X.,

RA Wang S. M., Woodage T., Wolley R., Sun B.,

RA Weilland S., Salden K. Bahler G., S., Chan M., Shong R., Shong R.,

RA Weilland R., Shong R., Rubin G. W., Von We
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10;
                                                                                                                                                                                                                      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Enthropoda; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 91;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match 6.6%; Score 89.5; DB 5; Length 1377; Best Local Similarity 18.8%; Pred. No. 11; Matches 51; Conservative 42; Mismatches 87; Indels 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1377 AA; 152801 MW; 40129F32B178DA84 CRC64;
                                     Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Flybase; FBgn0016641; PTP-ER.
Interpro; IPR000387; TYR_phosphatase.
Interpro; IPR000242; Tyr_prot_phphtase.
        Created)
                                                                                                                                                                                                     Drosophila melanogaster (Fruit fly).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam; PF00102; Y_phosphatase; 1. PRINTS; PR00700; PRTYPHPHTASE. SMART; SM0194; PTPC; 1.
01-MAY-2000 (TrEMBLrel. 13, 01-MAY-2000 (TrEMBLrel. 13, 01-DEC-2001 (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                               PIP-ER OR CG9856.
                                                                                                                             PTP-ER PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Hydrolase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
```

```
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,
                                                                                                                                                                                   401 TIDWLMYHRKONPYQVQPIHCSSTIQSSLDSDASLIPSLGDFELKSACSVDGGSKFGIGA 460
                                                                                                                                                                                                                                                                                                                                                                                                               146 NTQIRSVYRNPE------LNQC-----AGGAAMSKHLTNSAIDIWVPDLEIK 186
                                                                                      -----VHNITESH 340
                                                                                                                                               ---EVVRYQAYLQ-----SRLGNYLPPMSQLLTTARSWQACG 112
                                                                                                                                                                                                                                                                                     --HEPYQLPPEHLW-------GQIVPTLHLYQ------DLKSRGILPA 145
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 83 YQAYLQSRLGNYLPPMSQLL-TTARSWQACGHE-----PYQLPPEHLWGQIVPTLHLYQ 135
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          871 LESYLESYLSLALLQASHLLPKEATILRVTPHDIEPILPPFSSPESYL----IRAIHLYE 926
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DLKSRGI-LPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAID-----IWVPDL---EI 185
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             754 LLSSKIPLPITINYLSAFTESLYIDVQDSVSKRLETLQKDPATIPFSVVFSDQLFHDPLHP 813
23 PIPTNPQVSPI---KTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHATKQ- 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17 LVACSAPIPTNPQVS------51KTPSVLITKDKIGDHHTHE 52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ------HVGLQAHFETWLQMHHATKQEVVR 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Genome sequence of an obligate intracellular pathogen of humans:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ouery Match 6.5%; Score 88; DB 16; Length 1004;
Best Local Similarity 19.7%; Pred. No. 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  40; Mismatches 88; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1004 AA; 114401 MW; BB89266ED55F1496 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-NOV-1998 (TrEMBLrel. 08, Created)
01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
EXODEOXYRIBONUCLEASE V, GAMMA.
                                                                       292 PIKNSSOPSPLPKPKTPTIKSTKEKARSLDSAANESELSIV---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Chlamydia trachomatis.
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
NCBI_TaxID=813;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT; 1004 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        187 SQALYELQNRLCQYWLEHGENQNFGLGLYAT 217
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        521 TTVSTSSMNLL----QRRGSNHSLTLNLHSS 547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         981 CSRLSNDVLSQLLSLFINQDSQQN 1004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STRAIN=D/UW-3/CX;
MEDLINE=99000809; PubMed=9784136;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           186 KSQALYELQNRLCQYWLEHGENQN 209
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Chlamydia trachomatis.";
Science 282:754-759(1998).
EMBL; AE001334; AAC68244.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              52; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        53 HDESVS-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Complete proteome.
SEQUENCE 1004 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RECC OR CT640
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Davis R.W.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  084646
                                                                                                                                            79
                                                                                                                                                                                                                                                                                 113
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
O84646
ID 084646
ID D7 01
D7 
                                                                    QQ
                                                                                                                                                                                                     g
                                                                                                                                            οχ
                                                                                                                                                                                                                                                                          δŻ
                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        δŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qγ
```

```
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Kroph A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     75 ATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPPEHLWGQIVPTLHLY 134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -----VEHMWSKLMGTSALV 116
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               135 QDLKS-----RGI-----LPANT--QIRSVYRNPELNQCAGGAAMSKHLTNSAIDIW 179
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         72; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16 MLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESV-SHVGLQAHFETWLQMHH 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              37 LIISYEPTIPEAPQQAPALTN--LTLEELLSGLETNIHYEILEGHVG-----YLQVND 87
                                                                                                                                                                                                                                                                                                             "Phylogeny of Rodentia (Mammalia) inferred from the nuclear-encoded
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1773;
                                                                                                                                                   Spermophilus tridecemlineatus (Thirteen-lined ground squirrel). 
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; 
Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae; 
Spermophilus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6.3%; Score 86; DB 11; Length 407; 21.6%; Pred. No. 4.8; tive 30; Mismatches 43; Indels 7
                                                                                                                                                                                                                                                                                                                                                                                                 1 1
407 407
407 AA: 43838 MW; A5123E1D8DC36ECF CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
                                                                                   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
INTERPHOTORECEPTOR BINDING PROTEIN (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     01-JUL-1997 (TrEMBLrel. 04, Created)
01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
HYPOTHETICAL 57. I KDA PROTEIN.
RV3829C OR MTCY409.01 OR MTCY01A6.40.
Mycobacterium tuberculosis.
                                    407 AA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             536 AA.
                                                                                                                                                                                                                                                                                                                                               Mol. Phylogenet. Evol. 19:290-301(2001).
EMBL; AF297278; AAK62257.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              88 IPGQEVL----SKLGGFL------
                                                               01-DEC-2001 (TrEMBLrel. 19, Created)
                                    PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRT;
                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
MEDLINE-21240538; PubMed-11341810;
DeBry R.W., Sagel R.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MEDLINE=98295987; PubMed=9634230;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Local Similarity 21.6% Matches 40; Conservative
                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FROM N.A.
                                                                                                                                                                                                                         NCBI_TaxID=43179;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 180 -VPDL 183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    162 TLPEV 166
                                                                                  01-DEC-2001
01-DEC-2001
                                                                                                                                                                                                                                                                                                                           gene IRBP."
                                                                                                                                                                                                                                                                                                                                                                                  Receptor.
                                                                                                                                                                                                                                                                                                                                                                                                               NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                  NON_TER
                                                 0924J7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        007794;
                               0924J7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      007794
RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 12
                  0924J7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δλ
```

```
235 SETNSI 240
                                                                                                                                                                                                                                                           201 SQTLKI 206
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              NCBI_TaxID=287;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pseudomonas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9L868;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                091868
                                                                                                                                                                                                                                                                                                                                                                    09I2X5
                                                                                                                                                                                                                                                                                                                       RESULT 14
Q912X5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9L868
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                      QQ
                                                                                                                                                                      qq
                                              δy
                                                                                                                                 δλ
                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                                                                                                                                              11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A. Budga A.-M., Chaillus S., Hissler L., Stentz R., Champomier-Verges M., Budga A.-M., Zagorec M.; Alpert C.-A., Zagorec M.; Physical and genetic map of the Lactobacillus sakel 23K chromosome.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          75 ATKQEVVRYQAYLQSRLGNYLPPM-----SQLLTTARSWQACGHEPYQLPPEHLWGQI 127
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       282 ILTSPIVVSAIAPDVTINELIDPAVLPS-----EIRDRYLR-----IDHRG------SY 324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             123 LWGQIVPTLHLYQDLKSRGILPAN----TQIRSVYRNPELNQCAGGAAMSKHLTNSAIDI 178
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       370 -----SRGIVPADPTVVLQIPSLH-DPSLAPAGKQAA-----SAFAM 405
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          70 LQMHHATKQE---VVRYQAY----LQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPPEH 122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  30 VSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHF-------ETWLQ-MHH 74
                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                    11 TLISSMLVACSAP-IPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHVGLQAHFETW 69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
Lactobacillus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                325 LOMHFALAQPPAFAAPYQALNDPSMQASMGIFCTPEQ----VQQQWEDC------
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.; "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."; Nature 393:537-544(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                75;
                                                                                                                                                                                                                                                                                                                                                                                                                68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             y match 6.3%; Score 85.5; DB 2; Length 343; Local Similarity 22.4%; Pred. No. 4.3; nes 55; Conservative 22. ...
                                                                                                                                                                                                                                                                                                                                                                    Query Match 6.3%; Score 86; DB 16; Length 536; Best Local Similarity 24.0%; Pred. No. 6.9;
                                                                                                                                                                                                                                                                                                                                                                                                                49; Indels
                                                                                                                                   STRAIN-H37RV;
Parkhill J.;
Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; 297188; CAB10023.1; -.
TubercuList; Rv3829c; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases EMBL, AF401678; AAL00966.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         343 AA; 38417 MW; 0A4FB7FBC419F63E CRC64;
                                                                                                                                                                                                                                       InterPro; IPR000171; Bac_phytoene_dh.
Pfam, PF02032; Phytoene_dh.
Hypothetical protein; Complete proteome.
SEQUENCE 536 AA; 57073 MW; 59A3906E91619FAE CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  =
                                                                                                                                                                                                                                                                                                                                                                                                                22; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequent-DEC-2001 (TrEMBLrel. 19, Last anno
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DNA POLYMERASE I (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  <del>:</del> = :
                                                                                                                                                                                                                                                                                                                                                                                                                  44; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          343
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Lactobacillus sakei
                                                                                                                 SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NCBI_TaxID=1599;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         406 WFP 408
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 179 WVP 181
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          01-DEC-2001
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NON_TER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                093CV9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 13
093CV9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
```

g δ QQ Óλ

ŏ

```
SEQUENCE FROM N.A.
STRAIN-ATCC 15692 / PAO1;
MEDLINE-20437337; PubMed-10984043;
Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
Stover C.K., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
Sanith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
"Complete genome sequence of Pseudomonas aeruginosa PAO1, an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                220 -----LIGLLQQTKEIGCLPGMQEILEVF----PNVCGGNIGPFQNLANIDLDMQQSLSP 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -----RYQGREVSLQRLTYFSPTLAYQVNDELSVGLSVGFSHQAVALNEDFRAPNO--- 219
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         125 GQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAID----- 177
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       76 TKQEVVRYQAYLQS--RLGNYLP-----PMSQLLTTARSWQACG-HEPYQLPPEHLW 124
                                                                                                                                                                                                                                                                 128 VPTLHLYQDLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDIWVPDLEIKS 187
                                                                                                                                              105 VANLHEYONVOTDEAVYGKGAKRAVPTEP-----ALLYQHLVQ-------KA 144
                                                                                                                                                                                                                      188 QALYELQNRLCQYWLEHGENQNFGL------GLYATGAIHLDTQGFRKWGAQF 234
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16 MLVACSAPIPTNPQVSPIKIPSVLIIKDKIGDHHTHEHDESVSHVGLQAHFETWLQMHHA 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           178 -----IWVPDLEIKSQALYELQNRL---CQYWLEHGEN-QNFGLGLYAT--GAI 220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               271 SFNLGVLWEPTDWFAWGATYQSESRMRLKGKYRVDYGQGWQGFWTGVHKSLGGAI 325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6.2%; Score 84.5; DB 16; Length 532;
22.6%; Pred. No. 9.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL, AE004602; AAG05153.1; -.
Hypothetical protein; Complete proteome.
SEQUENCE 532 AA; 58988 MW; BOE704243FDB92BE CRC64;
59 PKAVFGVKRTQVAANRLGVHLAGVDFDLLLASYLLNTTNNSNDLG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         01-MAR-2001 (TrEMBLrel. 16, Created)
1-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-0CT-2001 (TrEMBLrel. 18, Last annotation update)
HYPOTHETICAL PROTEIN PA1764.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 543 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                31; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Matches 53; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       opportunistic pathogen.";
Nature 406:959-964(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Pseudomonas aeruginosa.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best Local Similarity
```

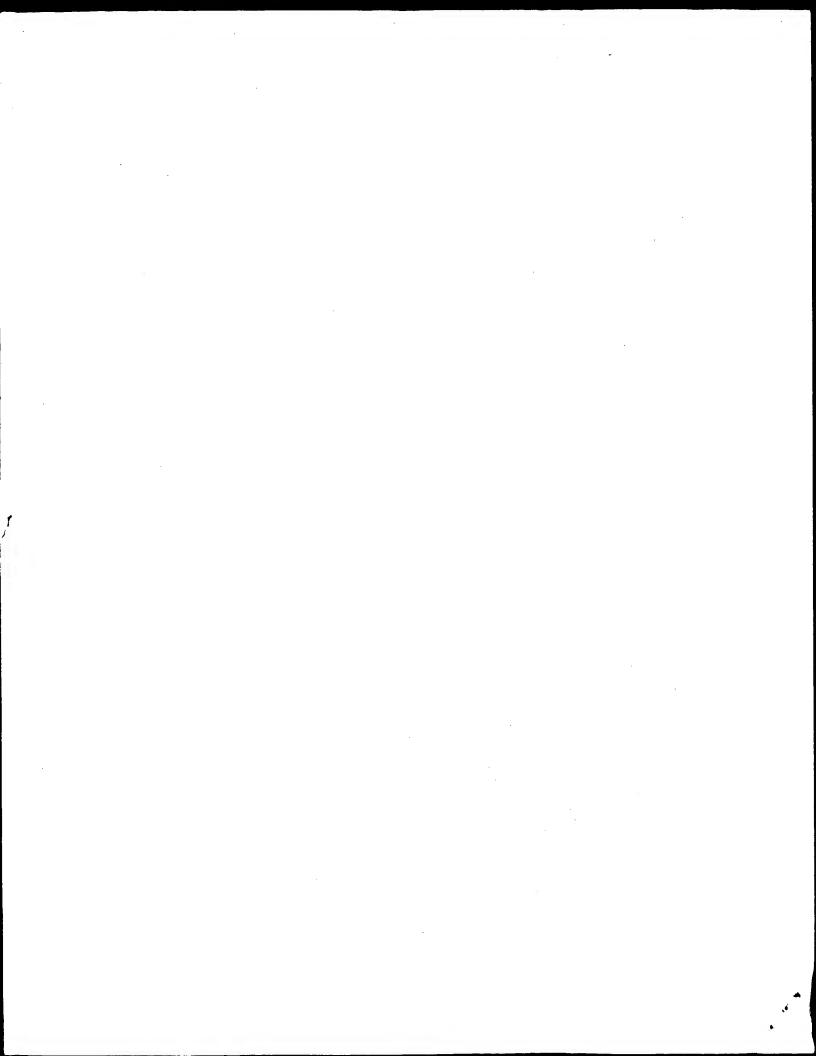
10;

```
PSOURDER FROM N.A.

RT Nucleotide sequence of the gene encoding the [NiFe] hydrogenase from RT Nucleotide sequence of the gene encoding the [NiFe] hydrogenase from RT Saralva L.M.;

RT Saralva
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | : ||| | : | : | : | : | : | 0 AINKPIPAN------ATYIRNLVLGNQFWHDHLVHFYHLHALDFVDVTSALLADPAK 140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  123 LWGQIVPTLHLYQDLKSRGILPANTQIRSVYRNPELN-QCAGGAAMSKHLTNSAI---- 176
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    19 ACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHV------GL 62
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
[NIFE] HYDROGERASE LARGE SUBUNIT (FRAGMENT).
Besulfovilbrio desulfuricans.
Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.
NCBI_TaxID=876;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match 6.2%; Score 84; DB 2; Length 543; Best Local Similarity 20.1%; Pred. No. 11; Matches 45; Conservative 31; Mismatches 90; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               177 ------DIWVPDLEIKSQALYELONRLCOYWLEHGENONF 210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       252 LYVKARAFIEEVXIPDLLLVASYYKD------WGKIGGTNNF 287
    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δy
```

Search completed: July 30, 2002, 16:19:11 Job time: 218 sec



DSK 1978

OM protein - protein search, using sw model

July 30, 2002, 16:15:58 ; Search time 29.76 Seconds

Run on:

(without alignments) 933.080 Million cell updates/sec

US-09-674-779-2

Perfect score: Sequence:

1 MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250

Gapop 60.0 , Gapext 60.0 OFIGO Scoring table:

747574 seqs, 111073796 residues

Searched:

Word size :

Total number of hits satisfying chosen parameters: Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1981.DAT:\*/SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1982.DAT:\* /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1980.DAT:\* A\_Geneseq\_032802:\*

/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/Aal983.Dar:
/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/Aal984.DAr:
/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/Aal986.DAr:
/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/Aal986.DAr:
/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/Aal986.DAr:
/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/Aal988.DAr:
/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/Aal988.DAr:
/SIDSI/gcgdata/hold-geneseq/geneseqp-embl/Aal988.DAr:

| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1992\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1993\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1993\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1995\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1995\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1997\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1999\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1999\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1999\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1999\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1999\_DAT:\*
| SIDSI/gcgdata/hold-geneseq/genesegp-embl/AA1999\_DAT:\* /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1990.DAT:\* .DAT:\* /SIDS1/gcgdata/hold-geneseq/geneseqp-embl/AA1991

/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\* Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Moraxella catarrha Description 250 22 AAB60645 Query Score Match Length DB 250 100.0 . 9

# ALIGNMENTS

AAB60645

AAB60645;

AAB60645 standard; Protein; 250 AA. X X ID

04-MAY-2001 (first entry)

Moraxella catarrhalis strain ATCC43617 BASB120 protein.

BASB120 protein; strain ATCC43617; antigen; antibody; vaccine; genetic immunisation; infection; upper respiratory tract; otitis media; hearing loss; deafness; pneumonia; sinusitis; nosocomial infection; invasive disease; antibacterial; auditory.

Moraxella catarrhalis

MO200109335-A2. 

Ş

08-FEB-2001

31-JUL-2000; 2000WO-EP07361.

99GB-0018281. 03-AUG-1999;

(SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.

Thonnard J;

WPI; 2001-159872/16. N-PSDB; AAF59797 New BASB120 polypeptides and polynucleotides from Moraxella catarrhalis strain American Type Culture Collection 43617, for use as therapeutic agents or vaccines against bacterial infections, e.g. otitis media or pneumonia

Claim 4; Page 64; 75pp; English.

BASB120 protein (AAB60645) and to DNA encoding it (AAF59797). The invention also relates to time who morability it (AAF59797). The invention also relates to immunogenic fragments of the BASB120 protein. Compositions comprising BASB120 nucleic acids, the recombinant production of BASB120, vaccine compositions comprising the anti-BASB120, and a method of BASB120 protein or nucleic acid, an antibody against BASB120, therapeutic compositions comprising the anti-BASB120 antibody, and a method of identifying a Moraxella catarrhalis infection via the detection of BASB120 proteins or antibodies. The vaccine compositions of the invention care useful as prophylactic or therapeutic agents against Moraxella catarrhalis infections in mammals, particularly humans. Moraxella catarrhalis in a Gram negative bacterium frequently isolated from the human upper respiratory tract, which is responsible for several pathological conditions. It is responsible for about 15% of cities media cases in children (which can lead to temporary or permanent hearing cases in children (which can lead to temporary or permanent hearing consocomial infections and, less frequently, invasive diseases. BASB120 proteins or nucleotides may additionally be used in screening for novel antibacterial compounds, and in the diagnosis and staging of infections.

The present sequence represents the Moraxella catarrhalis strain The invention relates to the Moraxella catarrhalis strain ATCC43617 ATCC43617 BASB120 protein. 

250 AA; Sequence

Gaps 1 MKNFNQYFITTLISSMLVACSAPIPTNPQVSPIKTPSVLITKDKIGDHHTHEHDESVSHV 60 0; Length 250; Indels 100.0%; Score 250; DB 22; 100.0%; Pred. No. 1.3e-258; tive 0; Mismatches 0; Best Local Similarity 100.( Matches 250; Conservative Query Match

δλ Q

ö

61 GLQAHFETWLQMHHATKQEVVRYQAYLQSRLGNYLPPMSQLLTTARSWQACGHEPYQLPP 120 ŏ

Оp

121 EHLWGGIVPTLHLYODLKSRGILPANTQIRSVYRNPELNQCAGGAAMSKHLTNSAIDIWV 180 ò

- 121 ehlwggivptlhlygdlksrgilpantgirsvyrnpelngcaggaamskhltnsaidiwv 180 엄
  - pp δy
- 241 CRHVLPKNKL 250 ||||||||||||||| 241 crhvlpknkl 250 δy

Search completed: July 30, 2002, 16:19:45 Job time: 227 sec

GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

July 30, 2002, 16:17:53 ; Search time 13.06 Seconds (without alignments) 467.565 Million cell updates/sec

US-09-674-779-2 250 Title: Perfect score: Sequence:

1 MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250

Scoring table:

Gapop 60.0 , Gapext 60.0

231628 seqs, 24425594 residues Searched:

Word size :

0 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Π Query Score Match Length DB Result

Description

No matches found

Search completed: July 30, 2002, 16:20:05 Job time: 132 sec

	·	
		•

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model Run on:

July 30, 2002, 16:18:23 ; Search time 19.33 Seconds (without alignments) 1242.749 Million cell updates/sec

US-09-674-779-2 250 Title: Perfect score:

1 MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250

Scoring table:

Sequence:

OLIGO Gapop 60.0 , Gapext 60.0

283138 seqs, 96089334 residues Searched:

Word size :

0 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

PIR\_71:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description Result Query
No. Score Match Length DB ID

No matches found

Search completed: July 30, 2002, 16:20:31 Job time: 128 sec

	······································		
	•		
			o
h			•

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
```

OM protein - protein search, using sw model

July 30, 2002, 16:20:08; Search time 10.97 Seconds (without alignments) 882.396 Million cell updates/sec Run on:

I MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250 US-09-674-779-2 250 Perfect score: Sequence:

Scoring table:

OLIGO Gapop 60.0 , Gapext 60.0

105224 seqs, 38719550 residues Searched:

Total number of hits satisfying chosen parameters: Word size :

0

Post-processing: Listing first 1000 summaries Minimum DB seq length: 0 Maximum DB seq length: 2000000000

SwissProt\_40:\* Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description Result Query
No. Score Match Length DB ID

No matches found

Search completed: July 30, 2002, 16:23:47 Job time: 219 sec

		,

```
GenCore version 4.5 Copyright (c) 1993 - 2000 Compugen Ltd.
```

Run on:

OM protein - protein search, using sw model

July 30, 2002, 16:19:48; Search time 26.55 Seconds (without alignments) 1628.954 Million cell updates/sec

US-09-674-779-2 250 1 MKNFNQYFITTLISSMLVAC......GAQFSETNSICRHVLPKNKL 250 Title: Perfect score:

Scoring table:

Sequence:

OLIGO Gapop 60.0 , Gapext 60.0

562222 seqs, 172994929 residues Searched:

Word size :

0 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

SPTREMBL\_19:\* Database :

1: sp\_archea:\*
2: sp\_bacteria:\*
3: sp\_fungi:\*
4: sp\_human:\*
5: sp\_invertebrate:\*
6: sp\_manmal:\*
7: sp\_mhc:\*
8: sp\_organelle:\*
9: sp\_phage:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

sp\_unclassified:\*

sp\_virus:\*
sp\_vertebrate:\*

sp\_rodent:\*

sp\_plant:\*

sp\_rvirus:\*
sp\_bacteriap:\*
sp\_archeap:\*

## SUMMARIES

Description Query Score Match Length DB Result

No matches found

Search completed: July 30, 2002, 16:23:30 Job time: 222 sec